THE INVERTED “U” RELATIONSHIP BETWEEN GOVERNMENT SIZE AND ECONOMIC GROWTH: DOES THE BARS CURVE APPLY TO PORTUGAL?

Sérgio Miguel Sousa Ferreira

Dissertation

Master in Economics

Supervised by

Maria Manuel Pinho

2022

Acknowledgements

At the end of this journey, I would like to express my gratitude to everyone that help and support me during this time, and without them this work would not be possible.

First, I would like to thank my supervisor, Professor Maria Manuel Pinho, for all the guidance and for always being available to help me, but also for the advice and motivation to improve my work and overcome the obstacles that appeared.

I would also like to thank my family, especially my parents and sister, for always believing in me and giving me all the support and encouragement to achieve my goals. I am also grateful to all my friends for their support, motivation and being there whenever I needed them.

Abstract: This dissertation aims to analyse the impact of general government expenditure and revenue on economic growth, by exploring if this relationship complies with an inverted "U" curve, as suggest by the BARS curve theory. To do so, besides the literature review on the subject, an empirical assessment for the Portuguese economy through ARDL models for the 1953-2019 period was conducted. The data is extracted from the new national accounts' dataset "Long time series for the Portuguese Economy" from INE and Banco de Portugal. The main results show an inverted "U" curve relationship between government size and economic growth in the long run, indicating a threshold of 23%-24% of GDP for the expenditure and of 23%-29% for the revenue. Furthermore, the relationship only holds for the current component of both expenditure and revenue. In general, the results are more robust for the expenditure side and suggest that in the long run, ceteris paribus, increases in government size induce a positive impact on economic growth only until a certain threshold.

JEL codes: C22, H10, H20

Keywords: government size, public expenditure, public revenue, economic growth, BARS curve

Resumo: Esta dissertação pretende analisar o impacto da despesa e da receita das Administrações Públicas no crescimento económico, explorando a conformidade com uma curva em "U" invertido, como sugere a teoria da curva BARS. Para tal, para além da revisão da literatura sobre o tema, foi realizada uma análise empírica para a economia portuguesa através de modelos ARDL para o período de 1953-2019. Os dados são extraídos do novo conjunto de dados de Contas Nacionais "Séries Longas para a Economia Portuguesa" do INE e do Banco de Portugal. Os principais resultados mostram uma relação de curva em "U" invertido entre a dimensão do governo e o crescimento económico no longo prazo, revelando um rácio ótimo de 23%-24% do PIB para a despesa e de 23%-29% para a receita. Ademais, esta relação apenas se revela significativa para a componente corrente, quer da despesa, quer da receita. Em geral, os resultados são mais robustos para a despesa e sugerem que, no longo prazo, ceteris paribus, aumentos na dimensão do governo induzem um impacto positivo no crescimento económico apenas até um certo ponto.

Códigos JEL: C22, H10, H20

Palavras-chave: dimensão do governo, despesa pública, receita pública, crescimento económico, curva BARS

Table of contents

List of figures

List of tables

1. Introduction

The intervention of governments in the economy and its relationship with economic performance has been a central topic of research for a long time. On the one hand, governments have an important role, not just by preventing a country from falling to a state of anarchy with no system of justice and discipline, but also on correcting market inefficiencies, enhancing macroeconomic stabilization and promoting equity. Therefore, government intervention becomes key in dealing with such issues since under certain conditions markets fail to be Pareto efficient, and even when they lead to efficient outcomes there may be a problem of inequality between individuals.

On the other hand, government intervention is not free from its own problems, namely when it deviates from the single goal of maximizing social welfare, and then creating government failures. When this situation overcomes the good side of government intervention, the effects can be contrary to the ones desired in the first place. In this context, the trade-off between solving market failures and creating government failures made new questions arise about the effects of the size of government on economic growth.

As such, many researchers tried to study a linear relationship between these variables, but it seems that there is no consensus on that approach. Some authors reached the conclusion of a positive effect, others of a negative effect, and others even an inconclusive result. So, a nonlinear hypothesis emerges. This new line of research, known as BARS curve, follows the seminal contributions of Barro (1989, 1990), Armey (1995), Rahn (Rahn & Fox, 1996) and Scully (1989, 1995, 2000, 2003), and shows an inverted "U" curve relationship between government size and economic growth.

Having this in mind, this dissertation will focus on studying this line of research. The purpose is to analyse the impact of government size on economic growth for Portugal. Therefore, we will be answering the following research questions: (i) is there a BARS curve in the Portuguese economy? (ii) and, at which size of the general government, is the economic output maximized in the last seven decades?

To answer these questions, we will use data from the new national accounts "Long time series for the Portuguese Economy" that were made available by Instituto Nacional de Estatística (INE) and Banco de Portugal on December 20, 2021. The analysis will cover the 1953-2019 period and will be conducted by a time series analysis through Autoregressive

Distributed Lag (ARDL) models, with two sub-analyses, one on the expenditure side and the other on the revenue side.

In this sense, the main motivations for this dissertation rest on several aspects. First, some global economic events have been intensifying the debate and research about government intervention and its effects on economic performance. For example, situations like the high unemployment caused by economic crises and the demographic aging problem in many developed countries can put the sustainability of the social security system at stake and as result of public finances as well. Moreover, the recent COVID-19 pandemic had serious impacts on the economies namely in employment and Gross Domestic Product (GDP), forcing governments to intervene in the economy with protection programs on households and firms to minimize the effects of the economic contraction, showing how key the government intervention is to attenuate such problems.

Beyond this, the literature review on the subject highlighted the research gap yet to be explored with most of the empirical studies suffering from the lack of harmonized data to measure the total government size and the lack of long observation periods. The last issue is indeed critical to access how the government size will impact the economic growth in the long run. Therefore, there seems to be a need for new evidence regarding the subject, namely for Portugal. In fact, this country has not only been recently stressed by the global economic events already mentioned, which make quite pertinent the study of this relationship, but also the release of a new long harmonized dataset will now allow to explore further this relationship for the country. Then, we aim to contribute to the scientific knowledge with more evidence on the BARS curve, first by analysing the impact of total government size on economic growth and, second by exploring the effects of government size components on economic growth for Portugal.

This dissertation will then start with the literature review of the seminal contributions of Barro, Armey, Rahn and Scully, exploring in detail the arguments presented by each author, the evolution of the arguments over time and the relevant empirical studies already made in some countries. Then, in section 3, we proceed to discuss how to put the theory into practice by describing the appropriate methodology to answer our research questions and performing an empirical analysis for Portugal. Finally, we draw the main conclusions and discuss the path ahead.

2. The relevant literature on the BARS curve

This chapter aims to explore the BARS curve literature, namely the seminal contributions and its empirical evidence. The BARS curve explains the inverted "U" curve relationship between government size and economic growth. The main idea is that there is a certain size of government at which economic growth is maximized. This means that government size, usually measured by the government expenditure or the government revenue (essentially, taxes) as a percentage of Gross Domestic Product (GDP), will have a positive impact on economic growth only until a certain threshold, and after that, general welfare starts to decrease.

Kuznets (1973, p. 247) defined economic growth in his Nobel prize lecture as "long-term rise in capacity to supply increasingly diverse economic goods to its population, this growing capacity based on advancing technology and the institutional and ideological adjustments that it demands". Although measuring the performance of the economy and the corresponding well-being of society is not easy, real GDP growth seems to be the most used measure. However, since the goal is trying to maximize the needs of the population, real GDP per capita growth is often used as a proxy to associate the measure of growth with social welfare.

2.1. Seminal contributions on the BARS curve

This section explores the theoretical framework on the inverted "U" curve relationship between government size and economic growth, following the seminal contributions of Barro, Armey, Rahn and Scully.

From Barro to Armey …

The first line of work on the BARS curve started with Robert J. Barro. After an initial work, in which the author tries to understand the determinants of per capita growth, he gets some interesting conclusions regarding government spending, namely that public investment (as a proxy for government infrastructure spending) tends to have a positive correlation with economic growth, unlike public consumption (Barro, 1989). This led to a new study on government spending, where the author creates a simple model of endogenous growth, analysing the effects of the government expenditure to GDP ratio on per capita economic growth. The main conclusions were two contrary effects: when the government size is small, an increase in the government expenditure ratio will lead to an increase in GDP per capita growth; and when the government size is too large, an increase in taxes will lead to a decrease in GDP per capita growth (Barro, 1990). Applying the model to the USA, the graphical result was an inverted "U" curve, where GDP per capita is maximized at a government expenditure ratio of 25% of the GDP.

If we look from the spending perspective, we may find some explanation for these results. As we already saw, the government needs to intervene in the economy, not only to provide a necessary legal and physical framework to prevent the country from falling to a state with no security or system of justice, but also to deal with market failures, including income disparities. However, besides the high costs of maintaining the proper framework, as we saw in Barro's argument, the spending that is made by the government is not all in productive goods and services. This means that there is a space to improve the framework and increase welfare by correcting market failures without seeing negative effects on GDP per capita, but only up to a certain point. After that, if spending continues to increase (which will also require more taxes to finance the spending raise), it will lead to a zone of negative returns.

After Barro (1989, 1990)'s contributions, a second line of work also emerged in the early 90s. With the problem of a high deficit being at the centre of the political debate in the USA, this new research on a nonlinear relationship between government size and economic growth not only got the attention of academics, but it also aroused the curiosity of politicians. The most popular one to get involved in this topic was Richard Armey, an US Congress Representative and economist, that argued that government intervention is needed to promote general welfare, but only up to some point (the so-called "optimal size"), and that beyond this threshold, government will become so large that general welfare will start to decrease (Armey, 1995).

Although Armey (1995) did not contribute with an empirical work (there is no estimation for this "optimal size"), his experience as politician and his economic knowledge allowed him to explore particularities of the nonlinear relationship, namely the effects of taxes on general welfare. To support his explanation, Armey based his work on the Laffer curve, popularized in the 70s by the economist and professor Arthur Laffer, that identified a trade-off between tax rates and tax revenues in his classes and meetings with political figures.

The Laffer curve shows the relationship between the tax rate and tax revenue, allowing for a revenue perspective on the government size. Laffer (2004) mentions two different effects that explain this inverted "U" curve relationship: the arithmetic and the economic effects. The first effect is based on the idea that more taxation will mean more revenue for the government to finance public goods and services and therefore better general welfare. However, according to the economic effect, this only works below some tax rate threshold since taxes have the particularity of creating distortions in the economy.

As pointed out by Laffer (2004, p. 3), tax revenue responds to tax rates according to "the tax system in place, the time period being considered, the ease of movement into underground activities, the level of tax rates already in place, the prevalence of legal and accounting-driven tax loopholes, and the proclivities of the productive factors". Therefore, raising taxes in situations where they are already too high can have the effect of penalizing the participation in the taxed activities, creating a negative impact on work and output, and consequently in the tax system that provides the revenue for the government (Laffer, 2004).

Several examples can explain this effect on the behaviour of taxpayers. For example, if we keep raising taxes on income, there will be less incentives for the private sector to work (since the government takes a high portion of the income), so taxpayers will probably move to leisure and work less, meaning less production and then less tax revenue. At the same time, the incentives to save will also decrease, which will mean less capital stock, leading to an additional negative impact on growth. If we take the example of indirect taxes, namely the increases in prices on products with low elasticity, it will mean that despite this increase consumers will not decrease the demand for those products, and therefore will create a higher burden for families with lower income. As regards increases in prices on products with high demand elasticity, the consumption of these products will tend to decrease, and thus the tax revenue as well. The same applies to the excess of taxation in firms, with the creation of a crowding out effect on private investment. But beyond that, taxpayers may be tempted to lower their tax liabilities, meaning "running" away from taxation by engaging in tax evasion (Stiglitz & Rosengard, 2015).

All these effects will mean that an increase in the tax rate in this zone when the tax rate is already too high could have an opposite effect in tax collection and fail to succeed in the goal of financing government expenditure. Therefore, an interesting point that the Laffer curve shows is that, after entering in this "prohibit range for government", as long as the tax rate keeps increasing, revenue will decrease (Wanniski, 1978).

Even if no empirical work was presented, the graphical result of this idea was again an inverted "U" curve, where government intervention will promote general welfare, but only up to some point, and that beyond this threshold, government will become so large that general welfare will start to decrease (the Armey curve). According to Armey (1995), the USA was at that time beyond this threshold.

In this sense, Armey (1995)'s work thus complemented Barro (1989, 1990)'s work. First, by relying on the Laffer curve and its explanations regarding the impacts of taxation on the economy, namely on economic growth. And second, by extending this analysis to general welfare.

and from Rahn to Scully …

The third seminal contribution was made by Richard Rahn. In his view, too little government is not favourable to economic growth, since without the public goods and services, markets will not function properly, namely on ensuring basic functions like policy protection, justice system, defence, property rights, and others. Nevertheless, the negative effects of the financial sources that allow the provision of public goods, such as "taxation, borrowing, or printing money", should not be disregarded (Rahn & Fox, 1996, p. 8).

According to Rahn and Fox (1996), taxation and borrowing represent a high cost on the share of national income, since both financing sources extract resources from the private sector of the economy. Moreover, it is possible to see similar conclusions to Armey (1995)'s work, by saying that individuals will not like transferring their income to the government, creating a negative impact on work and output (and consequently on economic growth). Having this in mind, it is only possible for government spending to increase income, if the public spending produces a higher increase in income than the reduction caused by the way that the spending is financed. Furthermore, his view on high government spending is also in line with Barro (1989, 1990)'s contribution, since he agrees that beyond some point, if the government size keeps increasing, "the greater the chances that it will be put to less-than-optimal use" (Rahn & Fox, 1996, p. 10). In addition to complementing the theory provided by the contributions of Barro (1989, 1990) and Armey (1995), the authors also provide an empirical analysis, reaching to the conclusion that, for the USA, the economic growth is maximized at a central government size of 15% of GDP.

The last contribution for the BARS curve was made by Gerald W. Scully. His work not only provides a contribution for the "optimal" size of government, but also suggests some interesting notes related with the efficient use of national resources. His initial work reached the conclusion that there is a negative correlation between the share of the size of government in the economy and economic efficiency (Scully, 1989). He explored this idea later and reached similar conclusions to the ones already seen in Barro (1989, 1990): when government spending as a share of GDP is low, as most of the public goods are productive (e.g. infrastructure, education, protection of property and others), an increase in taxes will mean a raise in economic growth; however, again this only works until a certain point by allowing to improve the provision of these productive goods and services, but once this point is passed, since the character of what is been providing is more non-productive spending (e.g. transfers and subsidies), an increase in taxes with this already large government will mean a lower economic growth rate (Scully, 1989, 1995). In this context, a new model was created, but now with a tax model perspective. Scully (2000, 2003) applied his model to New Zealand and the USA economies, also concluding for an inverted "U" curve, where the economic growth rate is maximized at the tax rate values of 20% of GDP and 19.3% of GDP, respectively.

The inverted "U" curve, known as the BARS curve, is then the result of years of work provided by these authors. The early works of Barro (1989, 1990), where the author studies the effects of productive and non-productive government spending on economic growth and later creates an endogenous growth model that explained this relation, continued with Armey (1995)'s contribution that, based on Laffer's work explored the effects of government size on general welfare. Then, the research proceeded with Rahn (1996) by developing a study that goes in line with the previous author, stressing the effects of financing sources, and also by providing a new empirical work on the issue; and finally with Scully (1989, 1995, 2000, 2003) and his new tax model and its applications to USA and New Zealand. Although these authors explored the issue in different ways, one thing was common in all the contributions: the result of an inverted "U" curve relationship between government size and economic growth.

2.2. Empirical evidence on the BARS curve

After Barro (1989, 1990), Armey (1995), Rahn (Rahn & Fox, 1996) and Scully (1989, 1995, 2000, 2003) have established the theoretical framework of an inverted "U" curve relationship between government size and economic growth, several empirical studies emerged in order to analyse the presence of the BARS curve in other countries. The studies were conducted by panel and individual country analysis through diverse empirical strategies, as shown in Table 1. This section explores the relevant empirical literature on the subject.

Lessons from panel data studies

One of the panel studies was conducted by Chobanov and Mladenova (2009), in which the authors analysed the relationship between general consumption expenditure and real GDP growth in 81 countries using information from the World Bank database. The use of final consumption expenditure instead of a government measure that covers total expenditure is due to the lack of availability of harmonized data for the latter, which leads many studies to choose to use only a part of the government spending. This can lead to biased results as the measure will not include expenditure such as public capital formation, penalizing the results by not addressing the effects of potentially more productive expenditure. Even so, it was possible to find evidence of an inverted "U" curve relationship, with economic growth being maximized at a general government consumption of 10.8% of GDP. These results were obtained through the estimation of a quadratic model using the Fixed Effects method. Robust standard errors were used to deal with heteroscedasticity and serial correlation issues.

Given the possibility of biased results, the empirical literature also tried to study the relationship with a better measure for government size and for a more restricted group of countries. Forte and Magazzino (2011) also studied the issue using a sample that only considers European Union countries. This will allow access to harmonized data covering the general government expenditure, which includes central government, local/regional government and social security funds. The study was conducted by a quadratic model and estimated by the GMM method to deal with possible endogeneity problems. Moreover, the authors recognize that the sample is composed of heterogeneous countries, so it became pertinent to subdivide the sample into more homogeneous groups. The division is made into four groups considering "the characters of their welfare and labour institutions, per capita GDP, and whether they have a common currency that determines their monetary policy"

(Forte & Magazzino, 2011, p. 311): Western Continental European countries, Anglo-Saxon countries, Mediterranean countries, and Eastern European countries.

The results again showed an inverted "U" curve relationship, with the peak of the curve varying between 38 and 42% of GDP depending on the subgroup under analysis and 37.79% of GDP for the EU as one group. Thus, the authors pointed out some interesting results: Western European countries group maximize economic growth at a lower point than Anglo-Saxon countries group, when the former is traditionally characterized by having more complex work institutions compared to the latter since Thatcher's reforms; the Mediterranean countries group obtained the bigger maximum point, and the Eastern European countries group also one of the lowest. However, it should be noted that, in the Eastern countries, the actual level of public expenditure on GDP is also one of the lowest, with the authors indicating the weak public economy structures, namely due to the "transition from the collectivist to the market economy system" (Forte & Magazzino, 2011, p. 312), as a possible explanation for these differences.

Beyond this, Barro (1989, 1990)'s seminal contribution has shown that not all government expenditure is productive, making it relevant to study the size of government expenditure components and how these productive components impact economic growth. So, a panel data analysis was conducted by Christie (2014), with data for 136 countries from the World Bank database. Christie (2014) studied the relation between central government expenditure and real GDP per capita growth. Though, in this study, the government size measure considers only central government expenditure, not taking into account the local/regional government and social security funds as in Forte and Magazzino (2011). However, it is still better than the use of government consumption as in Chobanov & Mladenova (2009). Since the data is broken down by functional classification, Christie (2014) tries to isolate productive elements of central government size, defining as "the sum of expenditure on education, health, housing, and transport and communication" (Christie, 2014, p. 186). The justification of the author for this sub-analysis relies on the effects of productive and non-productive spending being inseparable when only total government spending is considered.

In addition to this sub-analysis, Christie (2014) tried to make improvements in the empirical strategy to study this relationship. Therefore, besides an estimation of a linear and a quadratic model, a threshold model was also estimated. The latter model allows to address a non-linear relationship that is not only exactly quadratic as assumed by the quadratic model. The methodology then starts with testing the presence of an inflexion point with a threshold regression. After the threshold has been identified, its statistical significance needs to be verified. If it is statistically significant, the model is then estimated with standard econometric techniques with the threshold already founded, first by Fixed Effects and later with GMM. The GMM method is used to deal with possible endogeneity problems of government expenditure. Plus, an average data over 5 years interval is also used to attenuate cyclical effects changes and analyse the relationship in the long run.

The results of Christie (2014)'s work showed again an inverted "U" curve relationship with a threshold estimation of 33% for all countries. When the sample is divided in developed and developing countries, the results showed that economic growth is maximized at a government expenditure of 26% and 33% of GDP, respectively. However, the government expenditure coefficients for the quadratic model were never statistically significant and for the threshold model were not always statistically significant. Control variables (like inflation rate, investment, and openness to trade) were also used in the model and despite some exceptions their coefficients presented in general statistically significant results.

Moreover, Asimakopoulos and Karavias (2016) also studied the issue with a large World Bank sample and a segmented analysis for developing and developed countries. When compared with the studies already mentioned, one of the disadvantages of this study is the use of consumption as a measure of government size, instead of a total government expenditure measure. However, this study tries to bring an advantage over Christie (2014)'s work regarding the empirical strategy. Asimakopoulos and Karavias (2016) also use a threshold model to analyse this relationship, but instead of using the Hansen method used by Christie (2014), they used the Seo and Shin method for threshold regressions. In Christie (2014)'s strategy, first we need to find the inflexion point and later use that point to re-estimate again the model and deal with problems like endogeneity. The big advantage of the Seo and Shin method is that it already "allows for endogenous regressors and threshold variables" (Asimakopoulos & Karavias, 2016, p. 66), making it possible to deal with this problem from the beginning. Moreover, average data over a 5-year interval is again used to attenuate cyclical effects changes. Control variables were also used but not all of them were statically significant, and the authors pointed out that "the insignificance of some of the growth determinants maybe due to possible non-linear effects on growth instead of the linear assumed in our analysis" (Asimakopoulos & Karavias, 2016, p. 67). Results showed again that the developed countries seem to maximize economic growth at a lower point than the developing countries, as in Christie (2014)'s study. The larger share of non-productive spending in developed countries is pointed out as a possible explanation for these differences of 17.96% for developed countries and 19.12% of GDP for developing countries.

After exploring the empirical panel studies, it is possible to identify the main research complications. These issues could affect the results and therefore the thresholds that have been found by these empirical works. In this context, it is possible to identify three main problems.

(i) The lack of harmonized data to measure the total government size may lead to biased results

It seems that there is a lack of harmonized data to measure the total government size in a large period that is required for a panel analysis, making the researchers to only use a part of government size either in terms of expenditure components and government levels, which could lead to biased results. As already mentioned, final consumption can penalize the results by not addressing the effects of potentially more productive expenditure. Regarding the use of central government, this can also be a problem since it ignores expenditure made by regional and local governments. In countries where subnational governments have an important role through large public spending, this can be a problem. The only exception to this problem was the European Union study by Forte and Magazzino (2011) that use the general government expenditure; however, when splitting the sample in groups, they could not create a group for Scandinavian countries, since not all of them belong to the European Union and harmonized data was not available.

(ii) Modelling the relationship and controlling for institutional individual characteristics can be quite tricky

Every country has its own economic and political environment that can influence the size of the government. Therefore, it becomes a difficult task to model the relationship at hand and at same time controlling for different institutional characteristics of each country. For example, by just including different countries with different characteristics in the sample, the results can fail to explain the real relationship between the variables. Since this type of analysis only provides a threshold for the group as one, this could mean that the point where economic growth is maximized may not represent the real point for each country. Forte and

Magazzino (2011) divided the sample into homogeneous groups and obtained different results for each group, showing a good indication of this kind of problem. Moreover, economic growth may not depend only on the current value of government size as most of these studies assume, but also on its lagged values. However, to determine the appropriate lag length for a group of countries may not be an easy task for the reasons already mentioned and if we consider the possibility of the previous values of real GDP also being influenced the current value of real GDP, things can get even more tricky in choosing the right model structure and avoid endogeneity problems.

(iii) The size is not the only thing that matters

In all these studies, the majority of the variables used only take into consideration possible economic effects, ignoring political factors and the performance of public institutions, something that is clearly more difficult to measure given the subjectivity of the issue. For example, the size of the government is usually measured by the expenditure or the revenue as a percentage of GDP, but the regulation of the economy is also an important aspect of government intervention, although again not easy to measure. Moreover, political aspects like the ideology of the government can also be a factor of influence in the structure of the size of the government (Angelopoulos, Economides, & Kammas, 2012).

Besides this, the quality of the performance of the public institutions is also a key factor. As pointed out by Gwartney, Lawson, and Holcombe (1998, p. v), when government goes beyond its core function, negative effects happen on economic growth due to "(a) the disincentive effects of higher taxes, (b) diminishing returns as governments undertake activities for which they are ill-suited, and (c) an interference with the wealth creation process, because governments are not as good as markets at adjusting to changing circumstances and finding innovative new ways of increasing the value of resources". Furthermore, the efficiency losses due to rent seeking can also be a problem (Park, Philippopoulos, & Vassilatos, 2005), and even more difficult to measure.

Lessons from individual country studies

In this context, a set of time series analyses and econometric analyses with Instrumental Variables also emerged. As summarized in Table 1, many authors analysed individual countries, now accounting for the full size of public expenditure.

Going back to the Forte and Magazzino (2011)'s study, besides the panel data study, they also provide a time series analysis for the EU countries, estimating quadratic models through the ARIMAX method for the 1970-2009 period. To deal with possible problems of heteroscedasticity and autocorrelation, robust standard errors were used. The results showed a threshold range of 35.4% (Belgium) to 44.5% (Ireland) of GDP.

Other empirical studies have modelled the relationship using lagged GDP as an explanatory variable. As such, Altunc and Aydin (2013) perform a time series analysis. The authors decided to estimate quadratic models through the ARDL cointegration technique, which allowed to test a long-run relationship between government size and real GDP for Turkey, Romania, and Bulgaria for the 1995-2011 period. The results have shown the presence of an inverted "U" curve relationship, with real GDP being maximized at a general government expenditure of 25.21% (Turkey), 20.44% (Romania) and 22.45% (Bulgaria) of GDP.

Moreover, studying the relationship with a long period dataset is key to understanding this relationship. A study was made by Facchini and Melki (2013) considering a very long period and providing both a long and short-run analysis. As such, the authors analysed the relation between general government expenditure and real GDP and between general government expenditure and real GDP growth for France in the 1896-2008 period. Tests were also performed, and the results pointed to the existence of a cointegration relationship (Facchini & Melki, 2013); therefore, quadratic models were estimated with the FMOLS for the long-run analysis and next Error Correction Models were estimated for the short-run analysis. To take into account breaks in the time series, besides estimating the sample with time dummies, the authors split the series into two moments: the 1896-1938 period and 1947-2008 period. The results show the presence of an inverted "U" curve relationship in the long-run analysis, with the real GDP being maximized at a general government expenditure of 30% of GDP. Regarding the short-run analysis, the BARS curve was not observed when considering the full period of analysis, only when a shorter period (1947-2008) was used. Control variables like population and openness to trade were used, and despite some exceptions their coefficients presented in general statistically significant results.

Furthermore, Dobrescu (2015) also estimated quadratic models for Romania with a 1990-2013 period sample. The author performed the analysis between general government size and an annual index of real GDP with two different types of methods: a cointegration analysis through FMOLS, CCR and DOLS and an analysis using instrumental variables to address possible high correlation coefficient problems. Again, an inverted "U" curve relationship was found. Results showed that FMOLS and CCR obtained similar results with a maximum point of 32% and the DOLS methods obtained a maximum point of 29%. Regarding the estimation with instrumental variables, they obtained more statistically significant coefficients than the cointegration methods, with maximum thresholds of 32% with 2SLS and LIML and 31% with GMM.

When going back to the problems identified, it seems that the lack of harmonized measures for total government size has been solved. Regarding the second issue of the panel data studies only providing a threshold for the group, when each country is likely to have a different threshold given its individual institutional characteristics, the results of these individual country studies have shown different points at which economic growth is maximized, going in line with the literature that says that political and institutional particularities of each country may have an impact too. However, in the case of Romania, different studies concluded for different thresholds, which could also be explained by the different methodology and/or time samples used. Plus, the third issue remains, showing that finding variables that measure the quality of government can be quite difficult.

But beyond that, by looking at the data and methodology used in both panel and individual country analysis, we can get extra conclusions on the empirical assessments. It seems that there is a lack of consistency since some of them have interpreted the relationship as a short-run effect with annual growth rates and others as a long-run effect with 5-years average growth rates or through cointegration methods. Such divergence could also be an explanation for the different thresholds found by the studies. Furthermore, it seems that the empirical literature has focused on studying this relationship more from the expenditure side, and there still seems to be a gap in these empirical studies on the effects of the different expenditure or revenue components on economic growth.

Table 1: Empirical evidence on the BARS curve

Contributions	Sample	Type of analysis	Empirical strategy	Economic growth measure	Government size measure (% of GDP)	Control variables	Empirical threshold
Altunc & Aydin (2013)	1995-2011 (Turkey, Romania, and Bulgaria)	Individual countries study	ARDL (OLS)	Real GDP growth	General government expenditure	▪ Unemployment	**25.21%** (Turkey) **20.44%** (Romania) **22.45%** (Bulgaria)
Asimakopoulos & Karavias (2016)	1980-2009 (129 World Bank database countries)	Panel data study	GMM	Real GDP per capita growth	General government consumption	▪ Lagged real GDP per capita growth ▪ Gross capital formation (% of GDP) ▪ Inflation rate (consumer prices, %) ▪ Openness to trade (exports (E) + imports (I), as % of GDP) ▪ Population growth (%)	**18.04%** (all) **19.12%** (developing) **17.96%** (developed)
Chobanov & Mladenova (2009)	1961-2005 (81 World Bank database countries)	Panel data study	Panel Least Squares – fixed effects	Real GDP growth	General government consumption	-	**10.8%**
Christie (2014)	1971-2005 (136 World Bank database countries)	Panel data study	Two-way fixed effects; GMM	Real GDP per capita growth	Central government expenditure	▪ Inflation rate (consumer prices, %) ▪ Investment (% of GDP) ▪ Openness to trade (E+I as % of GDP) ▪ Average years of schooling ▪ Government effectiveness indicator	**26%-33%**

Contributions	Sample	Type of analysis	Empirical strategy	Economic growth measure	Government size measure (% of GDP)	Control variables	Empirical threshold
Dobrescu (2015)	1990-2013 (Romania)	Individual country study	FMOLS; CCR; DOLS; 2SLS; GMM; LIML	Annual index of real GDP	General government expenditure	▪ Gross fixed capital formation (as a geometric moving average index for two successive years)	**28.9%-31.9%**
Facchini & Melki (2013)	1896-2008 (France)	Individual country study	FMOLS	Real GDP growth	General government expenditure	▪ Openness to trade (E+I as % of GDP) ▪ Population size (total) ▪ Taxation (proportion of taxes in GDP) ▪ Dummies for World War periods	**30%**
Forte & Magazzino (2011)	1970-2009 (12 European Union countries)	Individual countries study	ARIMAX	Real GDP growth	General government expenditure	-	**35.39-44.47%**
	1970-2009 (27 European Union countries)	Panel data study	GMM				**37.79%** (all) **40.77%** (Anglo-Saxon) **39.85%** (Eastern European) **38.32%** (Western Continental European) **42.06%** (Mediterranean)

Source: Own elaboration.

Note: ARDL: Autoregressive Distributed Lag; OLS: Ordinary Least Squares; GMM: Generalized Method of Moments; FMOLS: Fully Modified Least Squares; CCR: Canonical Cointegrating Regression; DOLS: Dynamic Least Squares; 2SLS: Two-Stage Least Squares; LIML: Limited Information Maximum Likelihood; ARIMAX: Autoregressive Integrated Moving Average with Exogenous Variables.

3. Exploring the existence of a BARS curve for the Portuguese economy

This work intends to analyse the impact of the government size on Portugal's economic growth. In doing so, we will verify if this relationship complies indeed with a BARS curve as suggested by the literature review.

3.1. Data and empirical strategy

For the purpose already mentioned, it is fundamental to determine the main variables that will be considered in this empirical assessment. To account for the government size, general government data, which includes central, regional/local government and social security funds, will be used. Additionally, two sub-analyses will be made, one for the expenditure side and other for the revenue side. All fiscal variables will be considered as a percentage of GDP. Moreover, real GDP per capita (calculated at constant prices of 2016) is used to assess economic growth. By using a per capita measure, we seek to associate the measure of economic growth with social welfare. In addition, other variables that are expected to also impact economic growth will be used as control variables.

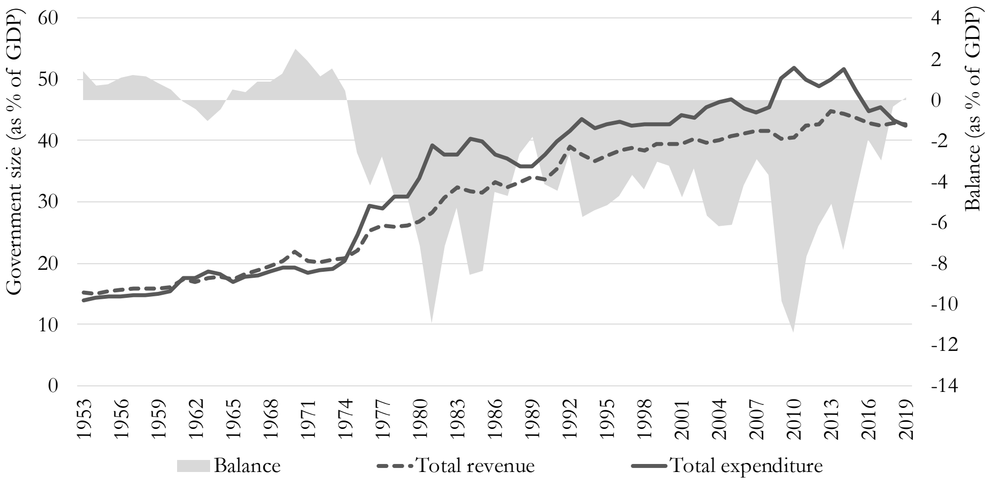
The data is collected from the new "Long time series for the Portuguese Economy" which were made available by Instituto Nacional de Estatística (INE) and Banco de Portugal on December 20, 2021. This dataset has a very long harmonized national accounts data for the 1947-2020 period. With 2020 being a particular year due to the COVID-19 pandemic and as the data is still provisional, this year will not be included in the analysis. Moreover, since GDP data is only available from 1953 onwards, the analysis will then cover the 1953-2019 period.

This section explores the dataset that comprises the main variables of interest and the appropriate methodology to be used in this study.

An overview of the relationship in the last decades

We start by looking at the general government size, namely general government expenditure and revenue. Figure 1 shows the evolution of these two variables in the last decades, as well as the corresponding balance. In this sense, it seems that in general there has been an increasing trend of both expenditure and revenue as a percentage of GDP, with the first having a superior increase compared to the second, which led to consecutive budget deficits, namely after 1974. However, since 2010 there seems to exist a decreasing tendency in the deficit, namely through the decrease in the expenditure-to-GDP ratio.

Figure 1: Total general government expenditure and revenue, 1953-2019



Source: INE and Banco de Portugal, "Long time series for the Portuguese Economy" (retrieved on March 10, 2022) and own calculations.

We can also further explore this dataset by looking at the components of these variables. To summarize such classifications, Table 2 displays the expenditure and revenue economic breakdown.

Table 2: General government revenue and expenditure economic classification

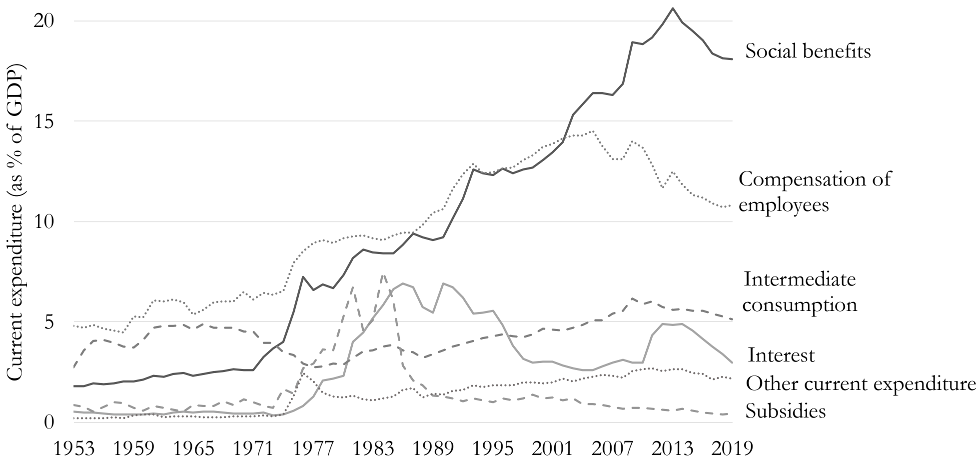
	Total expenditure	***Total revenue***
Current	▪ Social benefits ▪ Compensation of employees ▪ Interest ▪ Intermediate consumption ▪ Subsidies ▪ Other current expenditure	▪ Current taxes on income and wealth ▪ Taxes on production and imports ▪ Social contributions ▪ Sales ▪ Other current revenue
Capital	▪ Investment ▪ Other capital expenditure	▪ Capital revenue

Note: The descriptive statistics of the variables are available in Annex 1.

Having this classification in mind, we can now analyse these general government aggregates evolution. Figure 2 and Figure 3 show the current and capital expenditure as a percentage of GDP, respectively. Social benefits and compensation of employees are clearly the ones with the higher values as a percentage of GDP, with the first overtaking the number one place over the second around 2001. Plus, both variables show in general a tendency to increase over the period at hand (however, with a decrease in the last years).

This type of trend was not so notorious in intermediate consumption and other current expenditure that have experienced a slighter increase over the years. Subsidies had a significant increase around 1974 and a decline years later. Interests also had a significant increase after 1974 with a decline later, but a new increase happened again around 2009 with the global economic and financial crisis and later with the Portuguese need for international financial assistance (2011-2014). However, after such events, it is possible to observe a decline in interests.

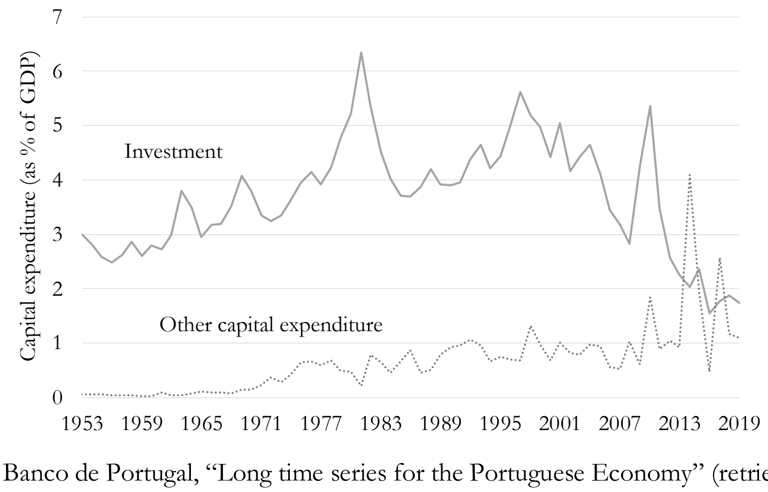
Figure 2: Current expenditure components, 1953-2019



Source: INE and Banco de Portugal, "Long time series for the Portuguese Economy" (retrieved on March 10, 2022) and own calculations.

Regarding capital expenditure (namely, investment), it seems to have had its ups and downs over the period at hand, but in general with a much lower value as a percentage of GDP when compared with the major components of current expenditure.

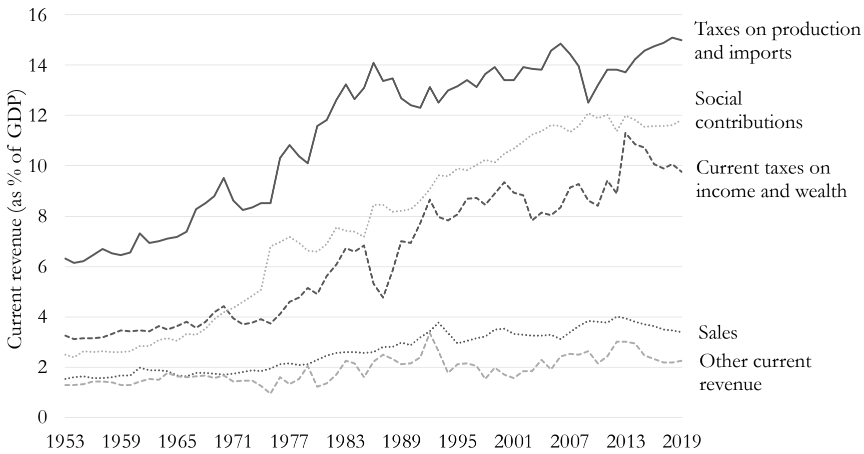
Figure 3: Capital expenditure components, 1953-2019



Source: INE and Banco de Portugal, "Long time series for the Portuguese Economy" (retrieved on March 10, 2022) and own calculations.

Regarding the general government revenue, Figure 4 and Figure 5 show the evolution of the current and capital components. Once again, the major categories (taxes and social contributions) show in general a tendency to increase, with taxes on production and imports being always the one with the highest share of GDP. Sales and other current revenue showed a slighter increase over the years than the already mentioned components.

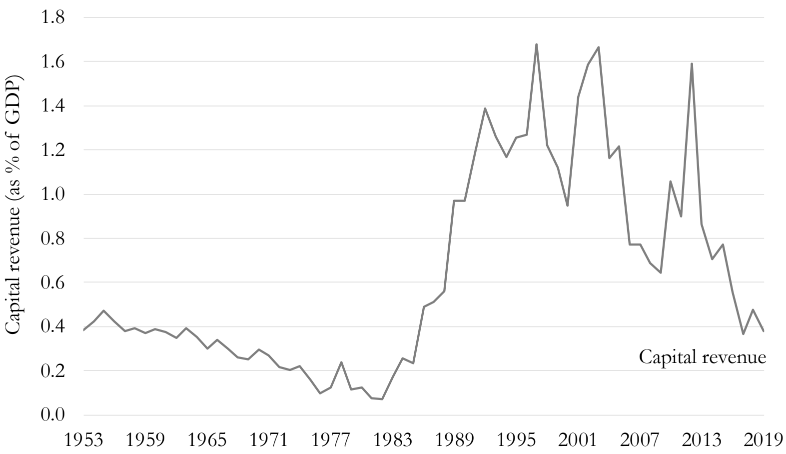
Figure 4: Current revenue components, 1953-2019



Source: INE and Banco de Portugal, "Long time series for the Portuguese Economy" (retrieved on March 10, 2022) and own calculations.

Moreover, capital revenue presented a decreasing trend until the 80s, followed by a significant increase with some ups and downs over the rest of the period. However, once again the capital revenue as a percentage of GDP was much lower when compared to the major current revenue components.

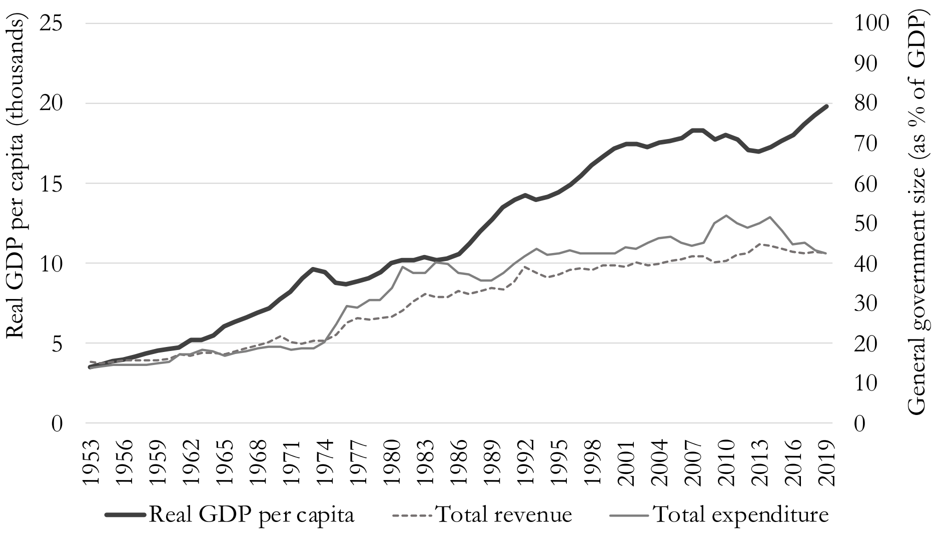
Figure 5: Capital revenue components, 1953-2019



Source: INE and Banco de Portugal, "Long time series for the Portuguese Economy" (retrieved on March 10, 2022) and own calculations.

Finally, Figure 6 presents the trajectory of real GDP per capita, alongside with the general government expenditure and revenue (in percentage of GDP). It is possible to observe that real GDP per capita series also shows in general a tendency of increasing over the period under analysis.

Figure 6: Real GDP per capita and general government size, 1953-2019



Source: INE and Banco de Portugal, "Long time series for the Portuguese Economy" (retrieved on March 10, 2022) and own calculations.

An introduction to the ARDL model

Several methods have been used to study this relationship over the years. However, one seems to stand out from the others: the empirical strategy based on an Autoregressive Distributed Lag (ARDL) model. Following the contribution of Pesaran, Smith, and Shin (2001), this type of model estimated by Ordinary Least Squares (OLS) has been used to analyse the relationship between variables in the long run, as well as to understand the implications of shocks that may occur in the short run.

The advantages over other methods can be summarized in three main points. First, as the name implies, it allows for both dependent and independent variables to be lagged. This can be quite useful, as it is not unlikely that past government size values affect economic growth in the current period, just as is not unlikely that past economic growth values affect the current economic growth period as well. With the help of information criteria (such as the Akaike information criterion), we can then define the appropriate lag structure for each variable, allowing for an adequate model estimation and therefore try to avoid endogeneity problems.

Moreover, considering that most series are non-stationary at levels, that is, they are characterized by the presence of trends, estimating such relationship can lead to spurious results by establishing a wrong relationship since the series could be growing over time maybe for other reasons and not due to the effect of one on the other. The common solution to solve such problems is to detrend the time series (Gujarati, 2012). This can be accomplished by two different processes. If a series has a determinist trend, we can just include one or more trend variables to model such tendency. But, if the trend is not deterministic, meaning the series has a stochastic trend and we are in the presence of a unit root process, we will have to differentiate the series in order for it to become stationary (Gujarati, 2012). The exception occurs when the variables are cointegrated (there is a long-run relationship between the two), meaning that there is a particular linear combination between them such that, although they are not stationary separately, they become stationary together. In this sense, the second advantage emerges, since the ARDL method provides a test for cointegration (the ARDL bound testing approach) allowing to check for a long-run relationship. Then, the cointegration technique also allows for the use of variables that are I(0)[1], I(1) or both, while using more traditional tests requires the variables to have the same order of integration.

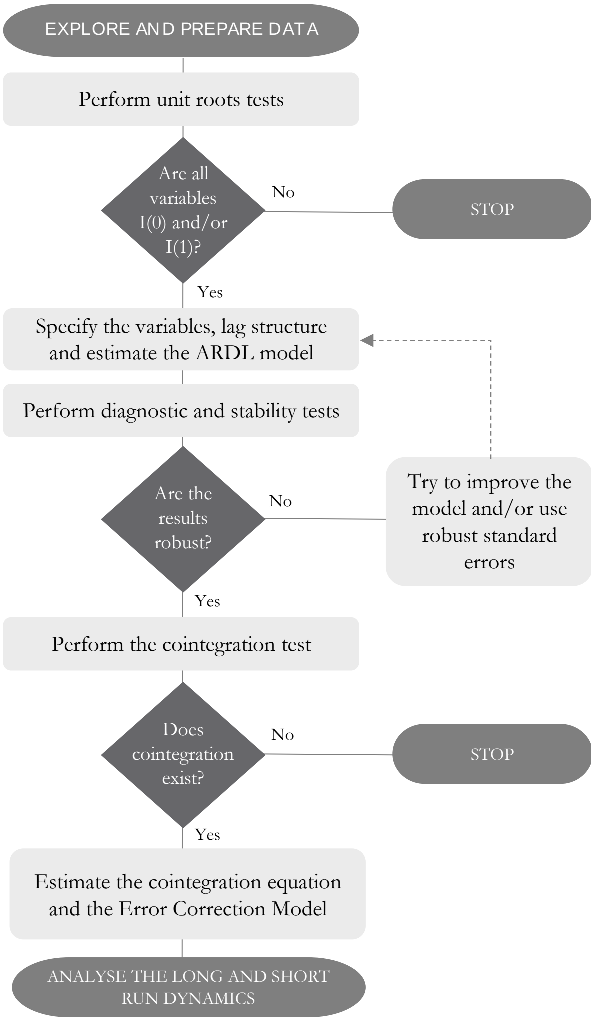
In addition, if indeed a cointegration relationship exists, we can use the long-run estimates to create a new variable (an error correction term) that will be added in the short-run analysis through an Error Correction Model (third advantage), considering long-run information together with the short-run effects. If a more traditional method that does not consider cointegration is used, with only differentiated variables in a short-run analysis, such approach will imply the loss of relevant long-run properties of this relationship.

Having this in mind, the choice will fall on the ARDL cointegration technique, since it seems to be the better way to study the relationship between government size and economic growth, when compared with other traditional methods used in the empirical literature. We recall that there is already one study that has used this method to study this relationship, the Altunc & Aydin (2013) study. Finally, in order to ensure the quality of the results, the study will be conducted through several tests and estimations before the final analysis. Figure 7 summarizes those steps. In the following section, each test will be explored in detail.

---

[1] A variable is said to be integrated of order d, I(d) if it becomes stationary after differencing it d times.

Figure 7: Empirical steps through ARDL



Source: Based on the chart of EViews Team (2017).

3.2. An empirical assessment through ARDL

As already mentioned, the purpose is to carry out the analysis of the impact of the government size on economic growth through ARDL models, estimated by OLS with the help of the statistical package EViews 12. This section will address in detail the above-mentioned stages of the method, which go from the preparation of the data until the final analysis of the results.

Exploring and preparing the data

After selecting the main variables and analysing them at levels in the previous section[2], we will now transform these variables into their natural logarithms. Such transformation will

[2] The descriptive statistics of the variables are available in Annex 1.

enable that when taking the first difference in the variables, the variables can be interpreted as an approximation to their growth rate.

In addition to economic growth (real GDP per capita) and government size (general expenditure and revenue) measures, squared government size variables were added to attempt to capture a possible quadratic relationship between government size and economic growth. Moreover, other variables such as openness to trade, labour force, government debt and a time dummy were used as control variables. With this, we now have nine main variables, as described in Table 3, with the respective descriptive statistics presented in Table 4.

Table 3: Variables of interest and the corresponding labels

Label	*Variable*
LN_REAL_GDPPC	Natural logarithm of real GDP per capita (REAL_GDPPC)
LN_EXP	Natural logarithm of general government expenditure as % of GDP (EXP)
LN_EXP_SQ	Natural logarithm of squared general government expenditure as % of GDP (EXP_SQ)
LN_REV	Natural logarithm of general government revenue as % of GDP (REV)
LN_REV_SQ	Natural logarithm of squared general government revenue as % of GDP (REV_SQ)
LN_OPEN	Natural logarithm of openness to trade [exports + imports, as % of GDP] (OPEN)
LN_LFORCE	Natural logarithm of labour force (LFORCE)
LN_DEBT	Natural logarithm of general government consolidated gross debt (DEBT)
DUMMY POST 1974	1 for the period 1974-2019, 0 otherwise

Table 4: Descriptive statistics of the variables of interest

Label	*Mean*	*Median*	*Maximum*	*Minimum*	*Std. Dev.*	*Observations*
LN_REAL_GDPPC	9.259	9.266	9.894	8.176	0.514	67
LN_EXP	3.424	3.631	3.949	2.632	0.444	67
LN_EXP_SQ	11.918	13.184	15.597	6.928	2.933	67
LN_REV	3.352	3.479	3.803	2.709	0.373	67
LN_REV_SQ	11.376	12.101	14.460	7.336	2.442	67
LN_OPEN	3.981	4.034	4.461	3.470	0.275	67
LN_LFORCE	8.335	8.375	8.547	8.045	0.181	67
LN_DEBT	8.623	9.680	12.429	3.956	3.111	67
DUMMY POST 1974	-	-	1	0	-	67

In the previous sub-chapter, we saw that exploring the time series through charts permits us to have some idea of their behaviour, namely if there is indication of a trend in the series. As such, Figure 8 and Figure 9 display new charts for the natural logarithms transformed series of the three main variables, in order to illustrate again the upward trend of each series, indicating that the series can be non-stationary.

Figure 8: Real GDP per capita (natural log scale), 1953-2019

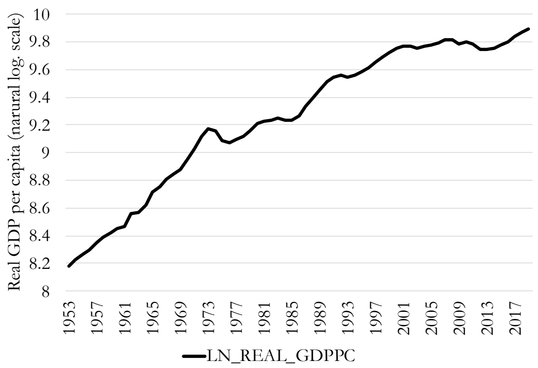
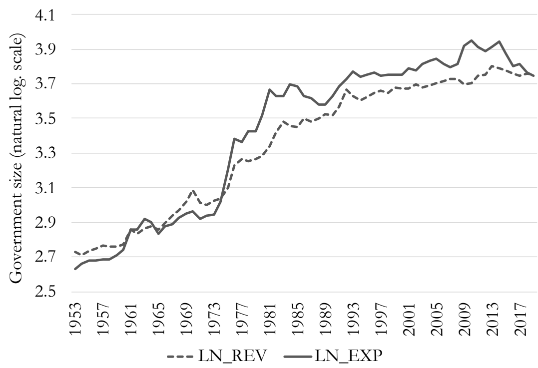


Figure 9: General government size (natural log scale), 1953-2019



Source: "Long time series for the Portuguese Economy" from INE and Banco de Portugal (retrieved on 10 March 2022) and own calculations.

Performing unit roots tests

Before estimating the ARDL models, it is important to check the stationarity of the series. Although in the presence of a cointegration relationship, the method enables the use of series that are individually non-stationary (since together they become stationary), we need to ensure that no series are I(2) or greater (not supported by the method), since in some cases the first differences is insufficient to obtain stationary series (Verbeek, 2004).

Thus, we performed unit roots tests, where the null hypothesis assumes that unit roots exist in the series. The existence of unit roots in the series means that we are in the presence of a non-stationary series. We conduct the tests both for levels and for the first difference to see if the null hypothesis is rejected for one of the cases for each variable. The results of the unit roots tests for the main variables are presented in Table 5.

Table 5: Unit roots tests

	Augmented Dickey-Fuller Test		***Phillips-Perron Test***	
	Level	First difference	Level	First difference
LN_REAL_GDPPC	0.9260	0.0005***	0.8615	0.0015***
LN_EXP	0.9576	0.0000***	0.9831	0.0000***
LN_EXP_SQ	0.9523	0.0000***	0.9817	0.0000***
LN_REV	0.9988	0.0001***	0.9900	0.0000***
LN_REV_SQ	0.9983	0.0002***	0.9824	0.0000***

Note: MacKinnon (1996) one-sided p-values; ***p-value < 0.01, **p-value <0.05, *p-value < 0.1; both tests include intercept and trend.

With the Augmented Dickey-Fuller and the Phillips-Perron unit roots tests, it is possible to conclude for all variables that the null hypothesis is rejected at first differences with a significance level of 1%. This means that all variables become stationary at first differences, leading us to rule out the possibility that there is a need for more differentiation. More tests for other variables in the dataset were performed (including tests for control variables) and like these results, most series are shown to be not stationary at levels but become stationary at first differences. The results are available in Annex 2.

Maximum lags and ARDL model selection

After inspecting the time series, the next step is to model the relationship. Following the empirical literature on the subject, a quadratic logarithmic model will be used in order to check for an inverted "U" curve relationship between the government size and economic growth. Two baseline specifications will be adopted considering only the main variables of the relationship, one for the expenditure side (E1) and the other for the revenue side (R1).

The method starts by analysing the equilibrium in the long run with the variables at levels, in our case at logarithmic levels. Equation (3.1.) illustrates the structure of the baseline specifications for the initial model.

$$Y_t = C + \sum_{i=1}^{k} \alpha_i Y_{t-i} + \sum_{i=0}^{q1} \lambda_i X_{t-i} + \sum_{i=0}^{q2} \sigma_i X^2_{t-i} + \varrho \, Trend + \varepsilon_t \qquad (3.1.)$$

Where: $\mathbf{Y_t}$ is the real GDP per capita (LN_REAL_GDPPC) at period t; $\mathbf{Y_{t-i}}$ is the real GDP per capita at period t-i; $\mathbf{X_{t-i}}$ is the general government expenditure or revenue as a percentage of GDP at period t-i (LN_EXP or LN_REV); $\mathbf{X^2_{t-i}}$ is the squared general government expenditure or revenue as a percentage of GDP at period t-i (LN_EXP_SQ or LN_REV_SQ); **Trend** is the linear deterministic trend term; $\boldsymbol{\varepsilon_t}$ is the white noise error term at period t; and **C**, $\boldsymbol{\alpha_i}$, $\boldsymbol{\lambda_i}$, $\boldsymbol{\sigma_i}$ and $\boldsymbol{\varrho}$ are the constant and coefficients of the other explanatory variables.

From here, we now must define the lags of the variables of our ARDL model (*k*, *q1*, *q2*), with *k* being the number of lags for the real GDP per capita, *q1* the number of lags for the general government expenditure or revenue and *q2* the number of lags of the squared general government expenditure or revenue.

Additionally, two more regressions will be estimated, one for the expenditure side (E2) and the other for the revenue side (R2). The additional models will have a structure similar to equation (3.1.), but the linear trend variable will be dropped from the baseline specifications and control variables will be included. The selected specifications were the ones that showed the best goodness of fit. Table 6 summarize these specifications.

Table 6: ARDL specifications

Specification	*Dependent variable*	*Explanatory Variables*
E1	LN_REAL_GDPPC	LN_EXP; LN_EXP_SQ; TREND
E2		LN_EXP; LN_EXP_SQ; LN_OPEN; LN_LFORCE; DUMMY POST 1974
R1		LN_REV; LN_REV_SQ; TREND
R2		LN_REV; LN_REV_SQ; LN_OPEN; LN_DEBT; DUMMY POST 1974

(i) Selecting the maximum lag structure

By using EViews 12 to estimate the ARDL model, we can use information criteria to choose the best lag structure for the model instead of trying to find manually the best lag structure by doing several estimations. However, we must point out the maximum number of lags we want. For that reason, we determine the maximum lag structure by the VAR Lag Order Selection Criteria for both the expenditure and revenue models, and the results are presented in Table 7 and Table 8, respectively.

Table 7: Maximum lag selection for the expenditure models

	Expenditure model				**Expenditure model (with control variables)**			
Lag	FPE	AIC	SC	HQ	FPE	AIC	SC	HQ
0	4.95e-05	-1.4002	-1.2981	-1.3600	1.80e-11	-10.5487	-10.2085	-10.4149
1	1.79e-09	-11.6267	**-11.2185***	**-11.4662***	1.94e-16	-21.9962	**-20.8056***	**-21.5279***
2	**1.67e-09***	**-11.7023***	-10.9880	-11.4214	**1.75e-16***	**-22.1183***	-20.0773	-21.3156
3	1.86e-09	-11.5958	-10.5753	-11.1945	1.91e-16	-22.0718	-19.1803	-20.9345
4	1.98e-09	-11.5442	-10.2175	-11.0224	2.48e-16	-21.8975	-18.1556	-20.4258

Note: * indicates the lag order selected by the criterion; FPE: Final prediction error; AIC: Akaike information criterion; SC: Schwarz information criterion; HQ: Hannan-Quinn information criterion.

Table 8: Maximum lag selection for the revenue models

	Revenue model				**Revenue model (with control variables)**			
Lag	FPE	AIC	SC	HQ	FPE	AIC	SC	HQ
0	1.03e-05	-2.9689	-2.8663	-2.9283	1.04e-09	-6.4931	-6.1529	-6.3593
1	3.78e-10	-13.1819	**-12.7737***	-13.0214	3.30e-15	-19.1595	**-17.9689***	**-18.6913***
2	**2.87e-10***	**-13.4593***	-12.7449	**-13.1783***	**2.95e-15***	**-19.2895***	-17.2484	-18.4867
3	3.15e-10	-13.3725	-12.3520	-12.9711	4.58e-15	-18.8973	-16.0057	-17.7600
4	3.72e-10	-13.2165	-11.8898	-12.6947	6.24e-15	-18.6711	-14.9291	-17.1993

Note: * indicates the lag order selected by the criterion; FPE: Final prediction error; AIC: Akaike information criterion; SC: Schwarz information criterion; HQ: Hannan-Quinn information criterion.

We selected a maximum of two lags for all specifications, as suggested by most of the tests and we only need to indicate a maximum of lags for EViews to define the best lag structure according to information criterion.

(ii) Determinate the lag structure

After selecting the maximum number of lags, we proceed to choose the best lag structure of the models according to the Akaike information criteria (AIC) as shown in Table 9.

Table 9: Selecting the lag structure through AIC

Specification	*Dependent variable*	*Explanatory Variables*	*1st choice by AIC*
E1	LN_REAL_GDPPC	LN_EXP; LN_EXP_SQ; TREND	ARDL (2,1,1)
E2		LN_EXP; LN_EXP_SQ; LN_OPEN; LN_LFORCE; DUMMY POST 1974	ARDL (1,1,1,2,1)
R1		LN_REV; LN_REV_SQ; TREND	ARDL (2,2,2)
R2		LN_REV; LN_REV_SQ; LN_OPEN; LN_DEBT; DUMMY POST 1974	ARDL (1,2,2,1,2)

Results show that for the baseline expenditure model, the selected structure is (2,1,1) and for the revenue model is (2,2,2). Regarding the additional models, the selected structure for the expenditure model is (1,1,1,2,1) and (1,2,2,1,2) for the expenditure and revenue model, respectively. The results of the different lags structures ordered according to the information criterion are presented in Annex 3.

Performing diagnostic and stability tests

With the model structure defined, we proceed with the estimation of the models. First, it is important to perform diagnostic and stability tests and understand if any improvement needs to be done, namely by adjusting the lags of the variables or using White/HAC robust standard errors. The null hypothesis (H_0) and the results of these tests are presented in Table 10.

Table 10: Diagnostic tests

	Breusch-Pagan-Godfrey ***Heteroskedasticity*** *test*	*Breusch-Godfrey* ***Serial Correlation*** *LM test*	*Jarque-Bera* ***Normality*** *test*
Null hypothesis	H_0: homoskedasticity	H_0: no serial correlation	H_0: normality
E1 (2,1,1)	1.0683 (0.3956)	0.2367 (0.7900)	0.5596 (0.7559)
E2 (1,1,1,2,1)	0.7443 (0.6917)	0.2147 (0.8075)	0.1181 (0.9426)
R1 (2,2,2)	**3.4437 ***** (0.0020)	1.2131 (0.3054)	1.0454 (0.5930)
R1 (2,1,1)	2.093 (0.0589)	0.0410 (0.9599)	1.4829 (0.4764)
R2 (1,2,2,1,2)	2.039 (0.0359)	0.8281 (0.4429)	0.2320 (0.8905)

Note: P-values in parentheses; ***p-value < 0.01.

The Breusch-Pagan-Godfrey test is used to check for heteroskedasticity. In all models, the null hypothesis of homoskedasticity is not rejected at 1% significance level, except the R1 model, in which the null hypothesis is rejected at 1% significance level, which leads to conclude that there is heteroskedasticity.

This is not a complete surprise since we already anticipated earlier problems like this in the design of the steps (see Figure 7). To solve this problem, and since this model also presented some instability in the CUSUM (Cumulative Sum) of Squares test (which verifies the stability of the coefficients over time), we adjust the lag structure for the second specification selected by the information criteria, the (2,1,1) structure. Now, the null hypothesis of homoskedasticity is not rejected at 1% significance level.

Concerning serial correlation, the Breusch-Godfrey LM test is used to check for the problem. In all the models, the null hypothesis of no serial correlation is not rejected at 5% significance level. Moreover, the Jarque-Bera test is used to check for normality, and again, in all the models, the null hypothesis of normality is not rejected at 5% significance level.

Finally, CUSUM and CUSUM of squares (CUSUMSQ) tests are used to check for stability of the coefficients over time by testing for structural breaks in the residuals. Figure 10, Figure 12, Figure 14 and Figure 16 show the results of the CUSUM tests, and Figure 11, Figure 13, Figure 15 and Figure 17 the CUSUMSQ results. Given that in all tests the lines do not exceed the limit bands for a significance level of 5%, it is concluded that the results are stable.

Figure 10: CUSUM (E1)

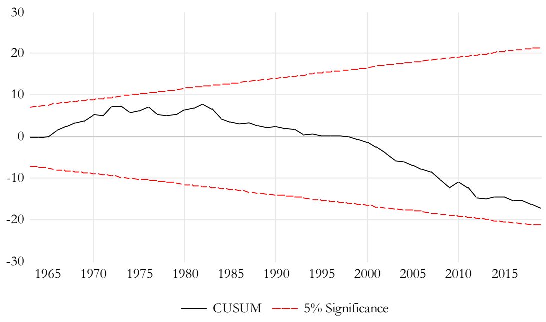


Figure 11: CUSUMSQ (E1)

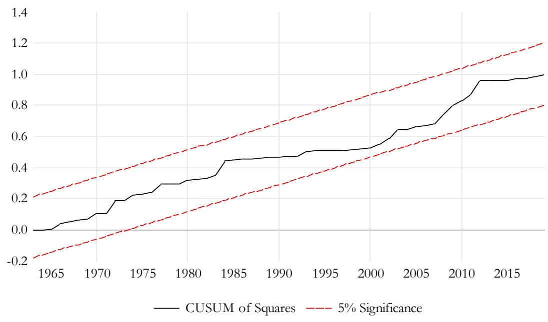


Figure 12: CUSUM (E2)

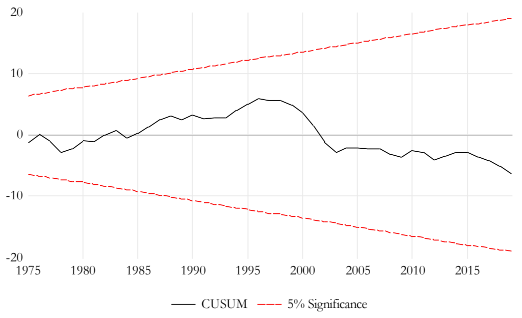


Figure 13: CUSUMSQ (E2)

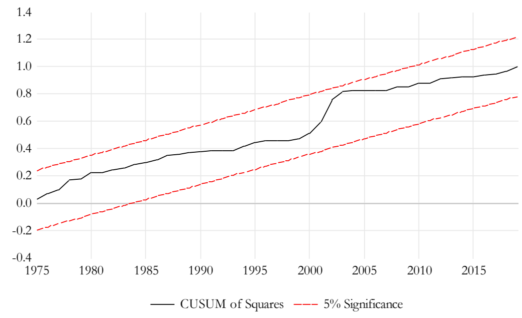


Figure 14: CUSUM (R1)

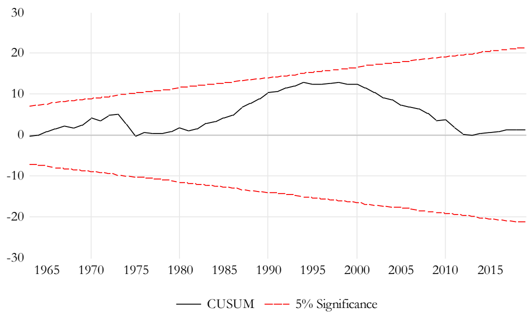


Figure 15: CUSUMSQ (R1)

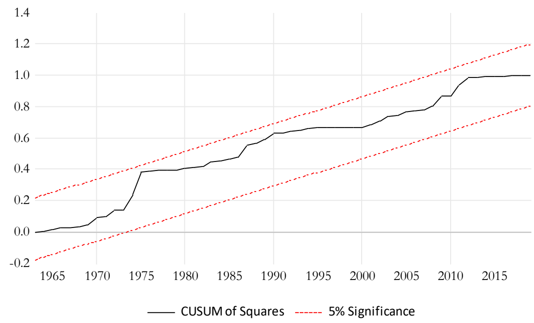


Figure 16: CUSUM (R2)

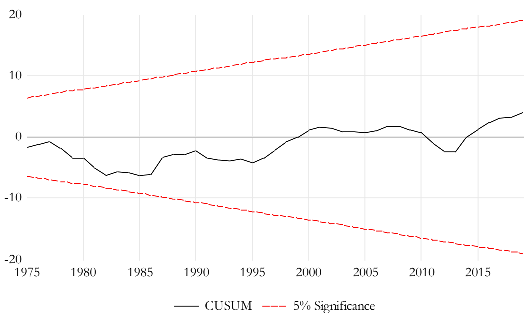
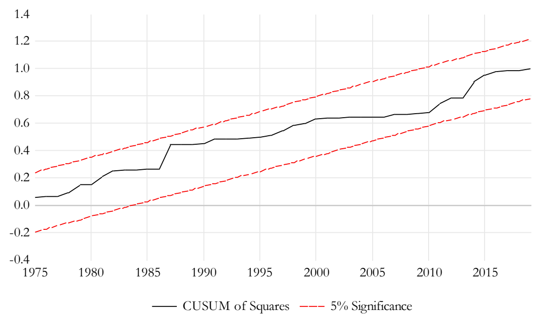


Figure 17: CUSUMSQ (R2)



After the ARLD models have passed the diagnostic and stability tests, we can then present the four models with their defined structure. Equation (3.2.), (3.3.), (3.4.) and (3.5.) represent these specifications.

$$LN\_REAL\_GDPPC_t = C + \sum_{i=1}^{2} \alpha_i \, LN\_REAL\_GDPPC_{t-i} + \sum_{i=0}^{1} \lambda_i \, LN\_EXP_{t-i} + \sum_{i=0}^{1} \sigma_i \, LN\_EXP\_SQ_{t-i} + \varrho \, Trend + \varepsilon_t \tag{3.2.}$$

$$LN\_REAL\_GDPPC_t = C + \sum_{i=1}^{1} \alpha_i \, LN\_REAL\_GDPPC_{t-i} + \sum_{i=0}^{1} \lambda_i \, LN\_EXP_{t-i} + \sum_{i=0}^{1} \sigma_i \, LN\_EXP\_SQ_{t-i} + \sum_{i=0}^{2} \theta_i \, LN\_OPEN_{t-i} + \sum_{i=0}^{1} \eta_i \, LN\_LFORCE_{t-i} + \gamma \, DUMMY\ POST\ 1974 + \varepsilon_t \tag{3.3.}$$

$$LN\_REAL\_GDPPC_t = C + \sum_{i=1}^{2} \alpha_i \, LN\_REAL\_GDPPC_{t-i} + \sum_{i=0}^{1} \lambda_i \, LN\_REV_{t-i} + \sum_{i=0}^{1} \sigma_i \, LN\_REV\_SQ_{t-i} + \varrho \, Trend + \varepsilon_t \tag{3.4.}$$

$$LN\_REAL\_GDPPC_t = C + \sum_{i=1}^{1} \alpha_i \, LN\_REAL\_GDPPC_{t-i} + \sum_{i=0}^{2} \lambda_i \, LN\_REV_{t-i} + \sum_{i=0}^{2} \sigma_i \, LN\_REV\_SQ_{t-i} + \sum_{i=0}^{1} \theta_i \, LN\_OPEN_{t-i} + \sum_{i=0}^{2} \tau_i \, LN\_DEBT_{t-i} + \gamma \, DUMMY\ POST\ 1974 + \varepsilon_t \tag{3.5.}$$

Performing the cointegration test

In the unit roots tests stage, the results give some information on the properties of the time series. Since they are non-stationary at levels even considering a deterministic linear trend, if a cointegration relationship does not exist, it means that the series have a stochastic trend. Therefore, estimating a regression at levels could lead to spurious results (Gujarati & Porter, 2009). If this situation is the reality, we should stop the process and rethink the empirical strategy and, for example, estimate the relationship through a model with the variables at first differences.

However, if cointegration exists between the time series, since they share a common stochastic trend (Verbeek, 2004), their "linear combination, so to speak, cancels out the stochastic trends in the two series" (Gujarati, 2012, p. 239). Thus, although separately the time series are non-stationary at levels, together they become stationary, with the error term now being I(0) (Verbeek, 2004).

Therefore, before proceeding with the analysis of the results, we need to check for cointegration in order to avoid spurious regressions. We use the ARDL Cointegration Bound Test (Pesaran et al., 2001) for that purpose. Annex 4 shows the criteria value to make the decision regarding the null hypothesis (no cointegration) and Table 11 presents the results of the tests. Thus, if the value of the test is superior to the upper limit, the null hypothesis is rejected (so, there is cointegration); if it is inferior to the lower limit, the null hypothesis is not rejected (so, there is no cointegration); and if it is between the two limits, the result is inconclusive.

Table 11: ARDL Cointegration Bound Test results

	No-cointegration		***Inconclusive***		***Cointegration***
E1 (2,1,1)		Lower limit I(0)		Upper limit I(1)	7.4685 (1%)
E2 (1,1,1,2,1)					4.4974 (2.5%)
R1 (2,1,1)					4.4295 (10%)
R2 (1,2,2,1,2)					9.8209 (1%)

For the expenditure models, the null hypothesis is rejected at a 1% (E1) and 2.5% (E2) significance level, concluding for the existence cointegration. Regarding the revenue models, the null hypothesis is also rejected at 10% (R1) and 1% (R2) significance level, also concluding for the existence of cointegration.

In addition, unit root tests were performed on the residuals of the regressions. Results show that the null hypothesis of having unit roots is rejected with a significance level of 1% for all the error terms, thus supporting the previous conclusion that says that together the series become stationary, with the error term being I(0).

Analysing the long-run equilibrium and the short-run dynamics

This sub-section explores the long-run equilibrium and short-run dynamics between the variables of interest.

(i) The long-run (equilibrium) relationship

Given the general ARDL regression, we can represent the long-run equation for the baseline specification (with a restricted time trend) as follows:

$$Y_t = C + \lambda_{LR} X_t + \sigma_{LR} X_t^2 + \varrho_{LR} \text{ Trend} + \varepsilon_t \qquad (3.6.)$$

where: $\mathbf{Y_t}$ is the real GDP per capita (LN_REAL_GDPPC) at period t; $\mathbf{X_t}$ is the general government expenditure or revenue as a percentage of GDP at period t (LN_EXP or LN_REV); $\mathbf{X_t^2}$ is the squared general government expenditure or revenue as a percentage of GDP at period t (LN_EXP_SQ or LN_REV_SQ); **Trend** is the linear deterministic trend term; $\boldsymbol{\varepsilon_t}$ is the white noise error term at period t; **C** is the constant; and $\boldsymbol{\lambda_{LR}}$, $\boldsymbol{\sigma_{LR}}$ and $\boldsymbol{\varrho_{LR}}$ are the long-run coefficients of the explanatory variables.

These long-run parameters of the baseline specification can be calculated based on (3.1.) as follows:

$$\lambda_{LP} = \frac{\sum_{i=0}^{q1} \lambda_i}{1-\sum_{i=1}^{k} \alpha_i} \quad (3.7.) \qquad \sigma_{LP} = \frac{\sum_{i=0}^{q2} \sigma_i}{1-\sum_{i=1}^{k} \alpha_i} \quad (3.8.) \qquad \varrho_{LP} = \frac{\varrho}{1-\sum_{i=1}^{k} \alpha_i} \quad (3.9.)$$

The additional models will have a structure similar to equation (3.6.), but the linear trend variable will be dropped from the baseline specifications and control variables will be included.

Therefore, these models will allow us to understand the long-run impact of the general government expenditure as a percentage of GDP (E1 and E2) or general government revenue as a percentage of GDP (R1 and R2) on real GDP per capita. Table 12 shows the long-run results.

Table 12: Long-run equilibrium – the levels equation

Dependent variable: *natural logarithm of real GDP per capita (LN_REAL_GDPPC)*							
Expenditure models				**Revenue models**			
Explanatory variables	E1	E2		R1	R2	***Explanatory variables***
LN_EXP	11.4831** (0.0383)	8.3302*** (0.0030)		7.3286*** (0.0062)	32.0624** (0.0145)	LN_REV
LN_EXP_SQ	-1.81240** (0.0494)	-1.3339*** (0.0023)		-1.0897** (0.0169)	-5.1143** (0.0237)	LN_REV_SQ
TREND	0.0270*** (0.0032)			0.0185** (0.0117)		TREND
LN_OPEN		1.3864** (0.0133)			0.4863 (0.2531)	LN_OPEN
LN_LFORCE		2.7521** (0.0367)			0.2350 (0.3993)	LN_DEBT
DUMMY POST 1974	-	YES		-	YES	DUMMY POST 1974
ARDL	(2,1,1)	(1,1,1,2,1)		(2,1,1)	(1,2,2,1,2)	***ARDL***
Akaike IC	-4.8895	-5.1624		-4.5503	-5.0162	Akaike IC
Schwarz IC	-4.6219	-4.7609		-4.2827	-4.5479	Schwarz IC
Threshold	24% of GDP	23% of GDP		29% of GDP	23% of GDP	**Threshold**

Note: P-values in parentheses; ***p-value < 0.01, **p-value <0.05 and *p-value < 0.1; Robust (White) standard errors and covariance to control for the possibility of heteroscedasticity.

Before analysing the results, we should recall that our models are quadratic logarithmic models. Thus, with a quadratic model, when analysing the effect of X on Y, we need to take into account the effect of X^2, since it does not make sense to hold X^2 fixed while X changes. Moreover, since Y and X (and X^2) are in natural logarithmic form, the combination of the coefficients X and X^2 represents the elasticity of Y with respect to government size.

Then, the impact of X on Y can then be interpreted through the following equation:

$$\frac{\Delta\%\hat{Y}}{\Delta\%X} \approx (\hat{\lambda} + 2\,\hat{\sigma}\,X)\% \tag{3.10.}$$

This means that, ceteris paribus, 1% change in X will induce approximately $(\hat{\lambda} + 2\,\hat{\sigma}\,X)\%$ change in Y, depending on the level of X.

Taking the results for the **E1 model**, in the long run, ceteris paribus, 1% change in general government expenditure (EXP) will induce approximately (11.5 - 2 x 1.8 EXP) % change in real GDP per capita, depending on the level of EXP. Regarding the **E2 model**, in the long run, ceteris paribus, 1% change in general government expenditure (EXP) will induce approximately (8.3 - 2 x 1.3 EXP) % change in real GDP per capita, depending on the level of EXP.

Now, taking the results for the **R1 model**, in the long-run, ceteris paribus, 1% change in general government revenue (REV) will induce approximately (7.3 - 2 x 1.1 REV) % change in real GDP per capita, depending on the level of REV. And concerning the **R2 model**, in the long run, ceteris paribus, 1% change in general government revenue (REV) will induce approximately (32.1 - 2 x 5.1 REV) % change in real GDP per capita, depending on the level of REV.

We should note that in all the models, the coefficients of both X and X^2 are statistically significant. Moreover, the negative value of the coefficient of X^2 confirms that indeed we have an inverted "U" curve relationship (BARS curve), where before the threshold (the turning point, X^*), X has a positive effect on Y, and after that point, X has a negative effect on Y. This threshold can be calculated by the following equation (3.11.).

$$X^* = \left| \frac{\hat{\lambda}}{2\hat{\sigma}} \right| \tag{3.11.}$$

Thus, in the **E1 model**, X^* will be approximately 3.168, which means that in the long run, real GDP per capita is maximized approximately at a general government expenditure of 24% ($e^{3.168}$ = 23.760) of GDP. Regarding the **E2 model** with control variables, it presents a similar threshold of 23% of GDP.

For the **R1 model**, X^* will be approximately 3.363, which means that, in the long run, real GDP per capita is maximized approximately at a general government revenue of 29% ($e^{3.363}$ = 28.876) of GDP. But concerning the **R2 model**, it presents a lower threshold (23%) than R1.

Furthermore, we should note that while in the E2 model all the coefficients of the control variables are statistically significant, in the R2 model the coefficients of the control variables are not statistically significant. Besides this, when comparing the thresholds of the revenue models, it is important to bear in mind that the structure of the R1 model is the second best chosen by the AIC (given that the first one did not pass the previous tests). Plus, the revenue models do not consider the same number of lags for the general government revenue variables, since the R1 model considers only up to 1 lag, while the R2 model takes into account up to 2 lags. So, some caution is needed when comparing the two revenue-side thresholds.

Finally, we can interpret the trend coefficient (ϱ_{LR}) of the baseline specifications, which, ceteris paribus, measures the change in Y from one year to the next (Wooldridge, 2013). Given that in the **E1 model**, ϱ_{LR}= 0.027, it means that, ceteris paribus, real GDP per capita grows about 2.7% per year on average (at a 1% significance level). Concerning the **R1 model**, where ϱ_{LR}= 0.0185, it means that, ceteris paribus, there is an approximate 1.85% increase in real GDP per capita per year on average (at a 5% significance level). This is in line with the existence of other factors not modelled but captured by the time trend that also affect real GDP per capita.

After performing the long-run analysis, we can use the long-run estimations to create error correction terms and add these new variables to new reparametrized models (Error Correction Models), which will consider both short and long-run information. For example, for the baseline specifications, the new variables can be represented by the following equations (3.12.) and (3.13.).

$$ECT_E = LN\_REAL\_GDPPC - (11.4831 \times LN\_EXP - 1.8124 \times LN\_EXP\_SQ + 0.0270 \times TREND) \quad (3.12.)$$

$$ECT_R = LN\_REAL\_GDPPC - (7.3286 \times LN\_REV - 1.0897 \times LN\_REV\_SQ + 0.0185 \times TREND) \quad (3.13.)$$

(ii) The short-run dynamics

In the previous sub-section, we analysed the long-run relationship, but short-run effects may occur deviating the variable from the long-run equilibrium and take some time to return to equilibrium, depending on the speed of adjustment. We can analyse these dynamics through an Error Correction Model (ECM). Equation (3.14.) represents the ECM for the baseline specifications.

$$\Delta Y_t = C + \sum_{i=1}^{n} \delta_i \, \Delta Y_{t-i} + \sum_{i=0}^{m1} \omega_i \, \Delta X_{t-i} + \sum_{i=0}^{m2} \psi_i \, \Delta X^2_{t-i} + \phi \, ECT_{t-1} + \mu_t \quad (3.14.)$$

where: Δ is the first difference of the variable; $\mathbf{Y_t}$ is the real GDP per capita (LN_REAL_GDPPC) at period t; $\mathbf{Y_{t-i}}$ is the real GDP per capita at period t-i; $\mathbf{X_{t-i}}$ is the general government expenditure or revenue as percentage of GDP at period t-i (LN_EXP or LN_REV); $\mathbf{X^2_{t-i}}$ is the squared general government expenditure or revenue as a percentage of GDP at period t-i (LN_EXP_SQ or LN_REV_SQ); $\mathbf{ECT_{t-1}}$ is the error correction term calculated from the long-run estimates (see 3.12 and 3.13.) at period t-1; μ_t is the white noise error term at period t; **C** is the constant; $\boldsymbol{\delta_i}$, $\boldsymbol{\omega_i}$, $\boldsymbol{\psi_i}$ and $\boldsymbol{\phi}$ are the coefficients of the other explanatory variables.

The additional models will have a structure similar to equation (3.14.), but control variables will be included.

Besides this, since the main variables are in their natural logarithmic form, their first difference can be interpreted as an approximation to their growth rate. For example, $\boldsymbol{\Delta Y_t}$ can be interpreted as an approximation to real GDP per capita growth.

Moreover, $\boldsymbol{\delta_i}$, $\boldsymbol{\omega_i}$ and $\boldsymbol{\psi_i}$ are the impact multipliers (the short-run effects), measuring the immediate impact that a change in the explanatory variable will have on a change in Y_t (Asteriou & Hall, 2007). Regarding $\boldsymbol{\phi}$, this coefficient is the feedback effect (or the adjustment effect), showing how much of the disequilibrium is being corrected, "i.e. the extent to which any disequilibrium in the previous period affects any adjustment in Y_t" (Asteriou & Hall, 2007, p. 310). Table 13 shows the results of the estimation of the ECM.

Table 13: Short-run dynamics – the Error Correction Model

Dependent variable: *real GDP per capita growth (D(LN_REAL_GDPPC))*					
Expenditure models			**Revenue models**		
Explanatory variables	E1	E2	R1	R2	***Explanatory variables***
C	-0.5431*** (0.0000)	-2.2315*** 0.0000	-0.3840*** (0.0001)	-2.7501*** (0.0000)	C
D(LN_REAL_ GDPPC(-1))	0.3016*** (0.0010)		0.4572*** (0.0000)		D(LN_REAL_ GDPPC(-1))
D(LN_EXP)	-1.4478*** (0.0016)	-0.8160** (0.0384)	-1.6029* (0.0698)	-1.3037* (0.0796)	D(LN_REV)
D(LN_EXP_ SQ)	0.1716*** (0.0096)	0.0871 (0.1293)	0.2200* (0.0995)	0.1658 (0.1414)	D(LN_REV_ SQ)
D(LN_OPEN)		0.0307 (0.3558)		1.6907** (0.0213)	D(LN_REV(-1))
D(LN_OPEN (-1))		-0.0835** (0.0128)		-0.2614** (0.0185)	D(LN_REV_SQ(-1))
D(LN_LFORCE)		1.6743*** (0.0000)		0.0949*** (0.0036)	D(LN_OPEN)
ECT(-1)	-0.0606*** (0.0000)	-0.0719*** (0.0000)		-0.1272*** (0.0001)	D(LN_DEBT)
				-0.0929*** 0.0065)	D(LN_DEBT(-1))
			-0.1138*** (0.0001)	-0.0639*** (0.0000)	ECT(-1)
DUMMY POST 1974	-	YES	-	YES	DUMMY POST 1974
ARDL	(2,1,1)	(1,1,1,2,1)	(2,1,1)	(1,2,2,1,2)	***ARDL***
Akaike IC	-4.9818	-5.2854	-4.6426	-5.1393	Akaike IC
Schwarz IC	-4.8146	-5.0178	-4.4754	-4.8048	Schwarz IC
F-statistic	25.3791	24.4246	13.7637	16.3477	F-statistic
Prob(F-statistic)	0.0000	0.0000	0.0000	0.0000	Prob(F-statistic)
R^2	0.63	0.75	0.48	0.73	R^2
Adjusted R^2	0.60	0.72	0.44	0.68	Adjusted R^2

Note: P-values in parentheses; ***p-value < 0.01, **p-value <0.05 and *p-value < 0.1.

Regarding the **E1 model**, ceteris paribus, if there is an increase of 1 percentage point (p.p.) in real GDP per capita growth rate in previous year, a positive change of 0.302 p.p. in the real GDP per capita growth rate at period t, at 1% significance level, is expected. Moreover, ceteris paribus, if there is an increase of 1 p.p. in general government expenditure (EXP) growth rate at period t, a change of approximately (-1.45 + 2 x 0.17 EXP) p.p. in the real GDP per capita growth rate at period t is expected.

Regarding the **R1 model**, ceteris paribus, if there is an increase of 1 p.p. in real GDP per capita growth rate in previous year, a positive change of 0.457 p.p. in the real GDP per capita growth rate at period t is expected, at 1% significance level. Moreover, ceteris paribus, if there is an increase of 1 p.p. in general government revenue (REV) growth rate at period t, one expects a change of approximately (-1.60 + 2 x 0.22 REV) p.p. in the real GDP per capita growth rate at period t.

We should note that in both **E1 and R1 models**, the coefficients of both X and X^2 are statistically significant, in the expenditure model at 1% significance level and in the revenue model at 10% significance level. However, an inverted "U" curve relationship is not observable, with the coefficient of X now being negative and of X^2 being positive in these models. The additional specifications (**E2 and R2**) presented similar results (except for some coefficients of the lagged variables of the R2 model). This situation could be explained by the fact that the variables in an ECM are in first differences of the natural logarithms, thus analysing the immediate impact of a variation in the growth of government size on economic growth.

We can also interpret the ECT(-1) coefficient (ϕ), which measures the adjustment to the long-run equilibrium and is expected to be negative so that the "the short run model converges to a long run solution" (Asteriou & Hall, 2007, p. 311). Given that in the **E1 model**, ϕ = -0.0606, it means that approximately 6% of the short-run discrepancy between the dependent and explanatory variables in the previous year is eliminated (converging back to long-run equilibrium) in the current period. Concerning the **R1 model**, where ϕ = -0.1138, it means that approximately 11% of the short-run discrepancy between the dependent and explanatory variables in the previous year is eliminated (converging back to long-run equilibrium) in the current period. Regarding the additional models **E2** and **R2** with control variables, the respective correction of such discrepancy is estimated to be approximately 7% and 6% per year.

Thus, such results indicate that there is an inverted "U" curve relationship in the long run between government size and economic growth series. Indeed, the results suggest that if we keep rising government size (measured by government expenditure or revenue) in the long run, it only has a positive impact on economic growth until a certain threshold. Such conclusion is in line with the literature review, which says that the type of spending (essentially, more productive or non-productive) and the burden for taxpayers may affect economic growth in such a way that it is only possible to increase general welfare directly through government size until a certain threshold.

The effects of the expenditure and revenue components

Despite having observed the presence of a BARS curve for Portugal, this does not mean that (all) the expenditure or revenue components comply with this inverted "U" curve effect. In this way, it becomes relevant to study the impact of such components on economic growth and thus try to fill the gap of BARS studies for such components. To this end, an analysis will be carried out with current government size isolated from capital government size. The main idea is to try to understand if indeed this inverted "U" effect comes from the current components.

The method (and respective steps) will be the same used in the previous analyses. As regards the specifications, since we do not know if all components show a BARS effect or if it is just de combination (total) of the components that give such effect, the first three specifications of the expenditure and revenue models will test linear and quadratic effects.

In addition to this, more specifications on the expenditure and revenue side will be made considering other variables that are expected to also impact economic growth. Thus, openness to trade, labour force, government debt and a time dummy will be used again as control variables. The selected specifications were the ones that showed the best goodness of fit.

Table 14 and Table 15 show the new variables of interest and the descriptive statistics, respectively.

Table 14: New variables of interest and the corresponding labels

Label	*Variable*
LN_REAL_GDPPC	Natural logarithm of real GDP per capita (REAL_GDPPC)
LN_CEXP	Natural logarithm of current expenditure as % of GDP (CEXP)
LN_CEXP_SQ	Natural logarithm of squared current expenditure as % of GDP (CEXP_SQ)
LN_KEXP	Natural logarithm of capital expenditure as % of GDP (KEXP)
LN_KEXP_SQ	Natural logarithm of squared capital expenditure as % of GDP (KEXP_SQ)
LN_REV	Natural logarithm of current revenue as % of GDP (CREV)
LN_REV_SQ	Natural logarithm of squared current revenue as % of GDP (CREV_SQ)
LN_CREV	Natural logarithm of capital revenue as % of GDP (KREV)
LN_KREV_SQ	Natural logarithm of squared capital revenue as % of GDP (KREV_SQ)

Table 15: Descriptive statistics of the new variables of interest

Label	*Mean*	*Median*	*Maximum*	*Minimum*	*Std. Dev.*	*Observations*
LN_REAL_GDPPC	9.259	9.266	9.894	8.176	0.514	67
LN_CEXP	3.268	3.491	3.845	2.384	0.487	67
LN_CEXP_SQ	10.914	12.189	14.780	5.685	3.064	67
LN_KEXP	1.430	1.472	1.974	0.701	0.283	67
LN_KEXP_SQ	2.125	2.168	3.896	0.491	0.789	67
LN_CREV	3.332	3.473	3.783	2.680	0.371	67
LN_CREV_SQ	11.238	12.062	14.312	7.181	2.412	67
LN_KREV	-0.737	-0.856	0.517	-2.605	0.807	67
LN_KREV_SQ	1.186	0.732	6.788	0.001	1.567	67

Table 16 and Table 17 show the results of government expenditure and revenue components models, respectively.

These results indicate that in the long run there is an inverted "U" relationship between government current accounts and economic growth. Then, suggesting that in the long run, ceteris paribus, increases on government current expenditure or revenue will induce a positive impact on economic growth only until a certain threshold. When considering current government expenditure, the results show thresholds of 19-20% of GDP. When considering current government revenue, the results show thresholds of 21-26% of GDP.

Regarding the capital government expenditure, which is a very volatile expenditure component, the coefficients are not statistically significant in general. However capital government revenue coefficients are statistically significant in some specifications that test the effect of this component as a linear rather than a quadratic effect. Moreover, while in the expenditure models all the coefficients of the control variables are statistically significant, in the revenue models not all the coefficients of the control variables are statistically significant.

Furthermore, in the short-run results, the inverted "U" curve relationship is only observable in some specifications, namely from the coefficients of the lagged current government accounts variables.

Table 16: Estimation results for expenditure components

	(1)	(2)	(3)	(4)
ARDL structure	(1,3,3)	(1,3,3,3,3)	(1,3,3,3)	(1,2,2,0,2,1)
Cointegration	YES [1%] 7.9607	YES [1%] 8.1421	YES [1%] 8.6568	YES [10%] 3.4105
Normality	YES (0.8875)	YES (0.2799)	YES (0.4304)	YES (0.6751)
No serial correlation	YES (0.3587)	YES (0.4341)	YES (0.2241)	YES (0.9504)
Homoscedasticity	YES (0.0796)	YES (0.1058)	YES (0.0790)	YES (0.3665)
CUSUM [5%]	ST	ST	ST	ST
CUSUMSQ [5%]	ST	ST	ST	ST
Akaike IC	-4.5562	-4.7636	-4.7494	-5.1567
Schwarz IC	-4.1851	-4.1227	-4.2434	-4.6549
Long run				
LN_CEXP	-6.5204 (0.9703)	8.2515* (0.0717)	8.3698** (0.0431)	6.6632** (0.0211)
LN_CEXP_SQ		-1.4154* (0.0642)	-1.4288** (0.0431)	-1.1130** (0.0137)
LN_KEXP	10.1057 (0.9688)	0.8696 (0.5734)	0.4237* (0.0657)	0.2016 (0.1632)
LN_KEXP_SQ		-0.1493 (0.7682)		
TREND	-0.4182 (0.9721)	0.0291*** (0.0003)	0.0287*** (0.0007)	
LN_OPEN				1.3048*** (0.0089)
LN_LFORCE				2.7737** (0.0252)
DUMMY POST 1974	-	-	-	YES
Current accounts threshold	-	19% of GDP	19% of GDP	20% of GDP

Table 17: Estimation results for revenue components

(5)	(6)	(7)	(8)	(9)	
(1,3,3)	(1,3,3,3,3)	(1,3,3,3)	(1,1,1,0,0, 2)	(1,1,1,0, 2)	***ARDL structure***
YES [2.5%] 5.4921	YES [1%] 9.4391	YES [1%] 10.1570	YES [1%] 9.0201	YES [1%] 10.6199	Cointegration
YES (0.7094)	YES (0.4671)	YES (0.8330)	YES (0.4453)	YES (0.3181)	Normality
YES (0.0131)	YES (0.3458)	YES (0.1034)	YES (0.3793)	YES (0.4080)	No serial correlation
YES (0.1190)	YES (0.7229)	YES (0.5708)	YES (0.0172)	YES (0.0148)	Homoscedasticity
NST	ST	ST	ST	ST	CUSUM [5%]
NST	ST	ST	ST	ST	CUSUMSQ [5%]
	-4.5895	-4.6045	-4.9451	-4.9570	Akaike IC
	-3.9486	-4.0985	-4.5437	-4.5890	Schwarz IC
		Long run			
	13.765*** (0.0000)	12.822*** (0.0000)	26.1436*** (0.0015)	26.9343*** (0.0009)	LN_CREV
	-2.1047*** (0.0000)	-1.9902*** (0.0000)	-4.2672*** (0.0035)	-4.4177*** (0.0022)	LN_CREV_SQ
	0.0278 (0.6892)	0.1254*** (0.0000)	-0.1343 (0.1637)	-0.1763** 0.0352	LN_KREV
	-0.0576 (0.1322)				LN_KEV_SQ
	0.0192*** (0.0004)	0.0227*** (0.0000)			TREND
			0.4216 (0.4198)		LN_OPEN
			0.3022 (0.1631)	0.3738* (0.0630)	LN_DEBT
-	-	-	YES	YES	DUMMY POST 1974
-	26% of GDP	25% of GDP	21% of GDP	21% of GDP	**Current accounts threshold**

Table 16: Estimation results for expenditure components

	(1)	(2)	(3)	(4)
ARDL structure	(1,3,3)	(1,3,3,3,3)	(1,3,3,3)	(1,2,2,0,2,1)
Short run				
C	0.0895*** (0.0000)	-0.3311*** (0.0000)	-0.3144*** (0.0000)	-2.0777*** (0.0000)
D(LN_CEXP)	-0.3143*** (0.0000)	-0.7552* (0.0598)	-0.6990* (0.0904)	-0.2306 (0.5102)
D(LN_CEXP_SQ)		0.0654 (0.2931)	0.0564 (0.3797)	0.0021 (0.9700)
D(LN_KEXP)	0.0140 (0.3781)	-0.1686* (0.0442)	0.0193 (0.1881)	
D(LN_KEXP_SQ)		0.0618** (0.0312)		
D(LN_CEXP(-1))	-0.0911 (0.1257)	0.2993 0.4839)	0.2521 (0.5620)	0.7905** (0.0186)
D(LN_CEXP_SQ(-1))		-0.0647 (0.3404)	-0.0555 (0.4194)	-0.1333** (0.0120)
D(LN_KEXP(-1))	0.0081 (0.6118)	-0.2404** (0.0187)	-0.0041 (0.7822)	
D(LN_KEXP_SQ(-1))		0.0817** (0.0189)		
D(LN_CEXP(-2))	-0.1395** (0.0162)	0.0350 (0.9343)	-0.1261 (0.7682)	
D(LN_CEXP_SQ(-2))		-0.0274 (0.6844)	-0.0032 (0.9625)	
D(LN_KEXP(-2))	0.0050 (0.7553)	-0.0604 (0.4443)	-0.0062 (0.6667)	
D(LN_KEXP_SQ(-2))		0.0132 (0.6287)		
LN_OPEN				0.0307 (0.3568)
LN_OPEN(-1)				-0.0796** (0.0139)
LN_LFORCE				1.5374*** (0.0000)

Table 17: Estimation results for revenue components

(5)	(6)	(7)	(8)	(9)	
(1,3,3)	(1,3,3,3,3)	(1,3,3,3)	(1,1,1,0,0, 2)	(1,1,1,0, 2)	***ARDL structure***
		Short run			
	-3.7610*** (0.0000)	-3.7180*** (0.0000)	-2.6672*** (0.0000)	-2.6554*** (0.0000)	C
	-1.1183 (0.1679)	-1.0102 (0.2033)	-1.3600* (0.0671)	-1.3361* (0.0749)	D(LN_CREV)
	0.1223 (0.3204)	0.0954 (0.4266)	0.1705 (0.1304)	0.1645 (0.1486)	D(LN_CREV_SQ)
	-0.0285** (0.0454)	-0.0025 (0.7817)			D(LN_KREV)
	-0.0105* (0.0508)				D(LN_KREV_SQ)
	-1.2846 (0.1819)	-1.2827 (0.1942)			D(LN_CREV(-1))
	0.1724 (0.2366)	0.1695 (0.2543)			D(LN_CREV_SQ(-1))
	-0.0447*** (0.0048)	-0.0514*** (0.0000)			D(LN_KREV(-1))
	0.0058 (0.2873)				D(LN_KREV_SQ(-1))
	-1.2280 (0.1528)	-1.2721 (0.1433)			D(LN_CREV(-2))
	0.1593 (0.2213)	0.1643 (0.2104)			D(LN_CREV_SQ(-2))
	-0.0332** (0.0280)	-0.0380*** (0.0010)			D(LN_KREV(-2))
	0.0025 (0.6393)				D(LN_KREV_SQ(-2))
			-0.1038*** (0.0011)	-0.1001*** (0.0016)	D(LN_DEBT)
			-0.1173*** 0.0008	-0.1283*** (0.0004)	D(LN_DEBT(-1))

Table 16: Estimation results for expenditure components

	(1)	(2)	(3)	(4)
ARDL structure	(1,3,3)	(1,3,3,3,3)	(1,3,3,3)	(1,2,2,0,2,1)
ECT(-1)	-0.0018*** (0.0000)	-0.0888*** (0.0000)	-0.0867*** (0.0000)	-0.0739*** (0.0000)
DUMMY POST 1974	-	-	-	YES
Akaike IC	-4.6499	-4.9199	-4.8744	-5.3105
Schwarz IC	-4.3801	-4.4476	-4.5033	-4.9760
F-statistic	9.2258	9.2392	10.3877	20.5415
Prob(F-statistic)	0.0000	0.0000	0.0000	0.0000
R-squared	0.54	0.71	0.66	0.77
Adjusted R-squared	0.48	0.63	0.60	0.73

Table 17: Estimation results for revenue components

(5)	(6)	(7)	(8)	(9)	
(1,3,3)	(1,3,3,3,3)	(1,3,3,3)	(1,1,1,0,0, 2)	(1,1,1,0, 2)	***ARDL structure***
	-0.2813*** (0.0000)	-0.3195*** (0.0000)	-0.0799*** (0.0000)	-0.0799*** 0.0000	ECT(-1)
-	-	-	YES	YES	*DUMMY POST 1974*
	-4.7457	-4.7295	-5.0989	-5.0800	Akaike IC
	-4.2735	-4.3584	-4.8648	-4.8459	Schwarz IC
	7.1480	8.2713	21.4447	20.8625	F-statistic
	0.0000	0.0000	0.0000	0.0000	Prob(F-statistic)
	0.65	0.61	0.69	0.68	R-squared
	0.56	0.54	0.66	0.65	Adjusted R-squared

Notes: LN_REAL_GDPPC is the depended variable in the long-run models; D(LN_REAL_GDPPC) is the depended variable in the short-run models; cointegration test: ARDL Cointegration Bound Test (Pesaran et al., 2001); normality test: Jarque-Bera test; serial correlation test: Breusch-Godfrey LM test; heteroskedasticity test: Breusch-Pagan-Godfrey test; stability tests: CUSUM and CUSUMSQ tests; ST = stable; NST = not stable; Given the instability of the model (5) coefficients (see CUSUM/CUSUMSQ test), the results are not presented; the lags of the variables in the specifications (1)-(3) and (6)-(7) were adjusted in order to guarantee the stability of the coefficients; Robust (White) standard errors and covariance to control for the possibility of heteroscedasticity; P-values are in parentheses; ***p-value < 0.01, **p-value <0.05 and *p-value < 0.1; ECT = Error Correction Term; IC = Information Criteria.

4. Conclusion

The inverted "U" curve relationship between government size and economic growth, also known as the BARS curve, is the result of years of work provided by Barro (1989, 1990), Armey (1995), Rahn (Rahn & Fox, 1996) and Scully (1989, 1995, 2000, 2003). Although these authors explored the issue in different ways, with some focusing more on the effects of the type of spending (essentially, more productive or non-productive) on economic growth, and others more on the tax burden as the main source of financing public expenditure, one thing was common in all the contributions: the result of an inverted "U" curve relationship between government size and economic growth.

Having this in mind, this dissertation aimed to analyse the impact of government size on Portugal's economic growth, through the ARDL cointegration technique, for the 1953-2019 period. The main results show an inverted "U" curve relationship between government size and economic growth in the long run, indicating thresholds of 23%-24% of GDP for the expenditure and of 23%-29% for the revenue. Furthermore, the relationship only holds for the current component of both expenditure and revenue, with thresholds of 19%-20% and 21%-26%, respectively. Thus, the results suggest that in the long run this relationship complies with a BARS curve and is more robust for the expenditure side.

It should be noted that this does not mean that short-run increases in government size will immediately give rise to a BARS effect. Indeed, in the short-run results, the inverted "U" curve relationship is only observable in some specifications. Moreover, as expected our long-run thresholds are different from those presented in other studies of individual countries, including the study for the countries of the European Union. As already mentioned, this is not unexpected given that, in addition to each country having its individual institutional characteristics, the studies do not use the same methodology nor time span or model the relationship in the same way (e.g., do not consider lagged variables or use control variables).

In the future, depending on data availability, the relationship can be further explored, namely by analysing the effects of the subcomponents of government current expenditure and revenue on economic growth and with more control variables that measure also the quality of the public institutions.

References

Altunc, O. F., & Aydin, C. (2013). The Relationship between Optimal Size of Government and Economic Growth: Empirical Evidence from Turkey, Romania and Bulgaria. *Logos Universality Mentality Education Novelty (Lumen 2013), 92*, 66-75. doi:10.1016/j.sbspro.2013.08.639

Angelopoulos, K., Economides, G., & Kammas, P. (2012). Does cabinet ideology matter for the structure of tax policies? *European Journal of Political Economy, 28*(4), 620-635. doi:10.1016/j.ejpoleco.2012.06.002

Armey, R. K. (1995). *The Freedom Revolution: The New Republican House Majority Leader Tells Why Big Government Failed, Why Freedom Works, and How We Will Rebuild America*. Washington, D.C.: Regnery Publishing.

Asimakopoulos, S., & Karavias, Y. (2016). The impact of government size on economic growth: A threshold analysis. *Economics Letters, 139*, 65-68. doi:10.1016/j.econlet.2015.12.010

Asteriou, D., & Hall, S. G. (2007). *Applied Econometrics: A modern Approach*: Palgrave Macmillan.

Barro, R. J. (1989). A Cross-Country Study of Growth, Saving, and Government. *National Bureau of Economic Research Working Paper Series, No. 2855*. doi:10.3386/w2855

Barro, R. J. (1990). Government Spending in a Simple-Model of Endogenous Growth. *Journal of Political Economy, 98*(5), S103-S125. doi:10.1086/261726

Chobanov, D., & Mladenova, A. (2009). What is the optimal size of government. *Institute for Market Economics (Bulgaria)*. Retrieved from https://www.researchgate.net/publication/280933948_What_is_the_Optimum_Size_of_Government

Christie, T. (2014). The Effect of Government Spending on Economic Growth: Testing the Non-Linear Hypothesis. *Bulletin of Economic Research, 66*(2), 183-204. doi:10.1111/j.1467-8586.2012.00438.x

Dobrescu, E. (2015). BARS Curve in Romanian Economy. *Amfiteatru Economic, 17*(39), 693-705. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2629594

EViews Team. (2017). AutoRegressive Distributed Lag (ARDL) Estimation. Retrieved from http://blog.eviews.com/2017/05/autoregressive-distributed-lag-ardl.html

Facchini, F., & Melki, M. (2013). Efficient government size: France in the 20th century. *European Journal of Political Economy, 31*, 1-14. doi:10.1016/j.ejpoleco.2013.03.002

Forte, F., & Magazzino, C. (2011). Optimal Size Government and Economic Growth in EU Countries. *Economia Politica, 28*(3), 295-321. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2333409

Gujarati, D. N. (2012). *Econometrics by Example*: Palgrave Macmillan.

Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics* (5th ed.): McGraw-Hill/Irwin.

Gwartney, J., Lawson, R., & Holcombe, R. (1998). The Size and Functions of Government and Economic Growth. *Joint Economic Committee*.

Kuznets, S. (1973). Modern Economic Growth: Findings and Reflections. *The American Economic Review, 63*(3), 247-258. Retrieved from http://www.jstor.org/stable/1914358

Laffer, A. (2004). *The Laffer Curve: Past, Present, and Future*. Backgrounder. The Heritage Foundation. Retrieved from https://www.heritage.org/taxes/report/the-laffer-curve-past-present-and-future

MacKinnon, J. G. (1996). Numerical Distribution Functions for Unit Root and Cointegration Tests. *Journal of Applied Econometrics, 11(6)*, 601–618. Retrieved from http://www.jstor.org/stable/2285154

Park, H., Philippopoulos, A., & Vassilatos, V. (2005). Choosing the size of the public sector under rent seeking from state coffers. *European Journal of Political Economy, 21*(4), 830-850. doi:10.1016/j.ejpoleco.2005.04.004

Pesaran, H., Smith, R., & Shin, Y. (2001). Bound Testing Approaches to the Analysis of Level Relationship. *Journal of Applied Econometrics, 16*, 289-326. doi:10.1002/jae.616

Rahn, R., & Fox, H. (1996). *What Is the Optimum Size of Government*. Vernon K. Krieble Foundation.

Scully, G. W. (1989). The Size of the State, Economic-Growth and the Efficient Utilization of National Resources. *Public Choice, 63*(2), 149-164. doi:10.1007/Bf00153397

Scully, G. W. (1995). The Growth Tax in the United-States. *Public Choice, 85*(1-2), 71-80. doi:10.1007/Bf01047902

Scully, G. W. (2000). The Optimal Size Fiscal State. In G. W. Scully & P. J. Caragata (Eds.), *Taxation and the Limits of Government* (1 ed., pp. 19-36): Springer.

Scully, G. W. (2003). Optimal taxation, economic growth and income inequality. *Public Choice, 115*(3), 299-312. doi:10.1023/A:1024223916710

Stiglitz, J., & Rosengard, J. (2015). *Economics of the Public Sector* (4th ed.). New York: W. W. Norton.

Verbeek, M. (2004). *A Guide to Modern Econometrics* (2nd ed.): John Wiley & Sons.

Wanniski, J. (1978). Taxes, Revenues, and Laffer Curve. *Public Interest, 50*, 3-16. Retrieved from https://www.nationalaffairs.com/public_interest/detail/taxes-revenues-and-the-laffer-curve

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach* (5th ed.): South-Western, Cengage Learning.

Annexes

Annex 1: Descriptive statistics

LABEL	VARIABLE	UNIT	MEAN	MEDIAN	MAXIMUM	MINIMUM	STD. DEV.	OBSERVATIONS
GDP	Gross Domestic Product (current prices)	10^6 euros	68 000.820	28 332.520	214 374.600	328.037	74 589.530	67
REAL_GDP	Real Gross Domestic Product (constant prices of 2016)	10^6 euros	117 438.400	106 082.500	203 854.900	30 498.870	56 888.060	67
POP	Population	10^3 residents	9 690.699	9 964.674	10 573.100	8 581.254	700.520	67
REAL_GDPPC	Real Gross Domestic Product per capita	euros	11 782.380	10 573.640	19 818.170	3 554.128	5 121.901	67
EXP	Total public expenditure (% of GDP)	%	33.501	37.748	51.899	13.903	12.702	67
CEXP	Current expenditure (% of GDP)	%	29.159	32.830	46.736	10.853	12.024	67
SOCIALB	Social benefits (% of GDP)	%	9.527	8.842	20.606	1.776	6.247	67
COMP_EMP	Compensation of employees (% of GDP)	%	9.578	9.430	14.523	4.486	3.231	67
INTEREST	Interest (% of GDP)	%	2.835	2.941	6.927	0.348	2.166	67
INT_CONS	Intermediate consumption (% of GDP)	%	4.324	4.274	6.162	2.727	0.874	67
SUBSIDIES	Subsidies (% of GDP)	%	1.527	0.980	7.481	0.390	1.555	67
OCEXP	Other current expenditure (% of GDP)	%	1.369	1.470	2.703	0.180	0.871	67
KEXP	Capital expenditure (% of GDP)	%	4.343	4.359	7.198	2.015	1.178	67
INVEST	Investment (% of GDP)	%	3.680	3.791	6.352	1.546	1.024	67
OKEXP	Other capital expenditure (% of GDP)	%	0.662	0.647	4.085	0.014	0.650	67
REV	Total public revenue (% of GDP)	%	30.457	32.417	44.817	15.007	10.228	67
CREV	Current revenue (% of GDP)	%	29.820	32.236	43.951	14.582	9.938	67
TAXES_INC	Current taxes on income and wealth (% of GDP)	%	6.400	6.581	11.299	3.132	2.541	67
TAXES_PROD	Taxes on production and imports (% of GDP)	%	11.258	12.621	15.080	6.142	2.994	67
SOCIALC	Social contributions (% of GDP)	%	7.597	8.158	12.088	2.408	3.432	67
SALES	Sales (% of GDP)	%	2.658	2.606	4.005	1.545	0.802	67
OCREV	Other current revenue (% of GDP)	%	1.907	1.788	3.350	0.965	0.514	67
KREV	Capital revenue (% of GDP)	%	0.637	0.425	1.677	0.074	0.456	67
BALANCE	Public balance (% of GDP)	%	-3.044	-3.216	2.477	-11.398	3.385	67
OPEN	Openness to trade (exports + imports, as % of GDP)	%	55.541	56.469	86.565	32.140	14.874	67
LFORCE	Labour force	10^3 residents	4 231.741	4 337.699	5 152.392	3 116.651	741.206	67
DEBT	General government consolidated gross debt	10^6 euros	59 766.710	15 998.160	249 977.500	52.230	81 493.400	67

Annex 2: Unit Root Tests

H_0: the series has a unit root

VARIABLE	*ADF TEST (based on AIC)*				*PP TEST*			
	Intercept		Intercept and linear trend		Intercept		Intercept and linear trend	
	Level	FD	Level	FD	Level	FD	Level	FD
LN_REAL_GDPPC	0.0414**	0.0008***	0.9260	0.0005***	0.0507*	0.0007***	0.8615	0.0015***
LN_EXP	0.4256	0.0000***	0.9576	0.0000***	0.3766	0.0000***	0.9831	0.0000***
LN_EXP_SQ	0.4790	0.0000***	0.9523	0.0000***	0.4502	0.0000***	0.9817	0.0000***
LN_CEXP	0.4653	0.0000***	0.9506	0.0000***	0.3725	0.0000***	0.9696	0.0000***
LN_CEXP_SQ	0.5325	0.0000***	0.9440	0.0000***	0.4841	0.0000***	0.9684	0.0000***
LN_KEXP	0.3318	0.0000***	0.9927	0.0000***	0.0366**	0.0000***	0.1926	0.0001***
LN_KEXP_SQ	0.2847	0.0000***	0.9602	0.0000***	0.0539*	0.0000***	0.0945*	0.0001***
LN_REV	0.2536	0.0001***	0.9988	0.0001***	0.4157	0.0000***	0.9900	0.0000***
LN_REV_SQ	0.3807	0.0001***	0.9983	0.0002***	0.5457	0.0000***	0.9824	0.0000***
LN_CREV	0.5221	0.0000***	0.9692	0.0001***	0.4084	0.0000***	0.9902	0.0000***
LN_CREV_SQ	0.6293	0.0000***	0.9506	0.0002***	0.5490	0.0000***	0.9706	0.0000***
LN_KREV	0.5244	0.0000***	0.7875	0.0000***	0.5186	0.0000***	0.7690	0.0000***
LN_KREV_SQ	0.2159	0.0162**	0.3795	0.0009***	0.2546	0.0000***	0.4281	0.0000***
LN_OPEN	0.7178	0.0000***	0.0295	0.0000***	0.8116	0.0000***	0.0211**	0.0000***
LN_LFORCE	0.5559	0.0067***	0.9527	0.0174**	0.6367	0.0117**	0.9809	0.0328**
LN_DEBT	0.4670	0.3022	0.8263	0.4895	0.7744	0.0528*	0.9590	0.1697

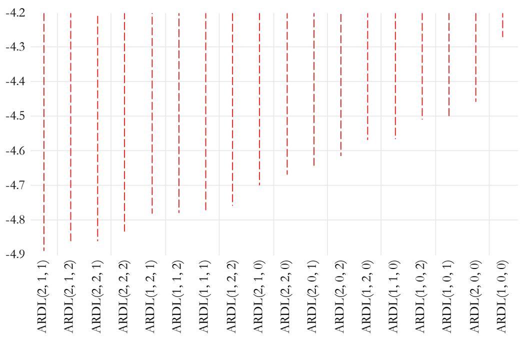
H_0: ε_t has a unit root

Specification	*ADF TEST*	*PP TEST*
E1	0.0000***	0.0000***
E2	0.0000***	0.0000***
R1	0.0000***	0.0000***
R2	0.0000***	0.0000***
1	0.0000***	0.0000***
2	0.0000***	0.0000***
3	0.0000***	0.0000***
4	0.0000***	0.0000***
5	0.0001***	0.0001***
6	0.0000***	0.0000***
7	0.0000***	0.0000***
8	0.0000***	0.0000***
9	0.0000***	0.0000***

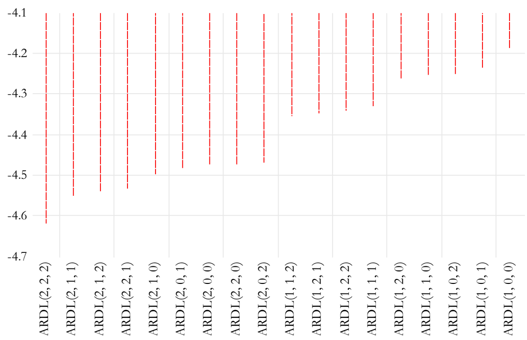
Notes: ADF: Augmented Dickey-Fuller Test; PP: Phillips-Perron Test; AIC: Akaike information criterion; FD: First Differences; MacKinnon (1996) one-sided p-values in the tables; ***p-value < 0.01, **p-value <0.05, *p-value < 0.1.

Annex 3: Determining the lag structure through AIC

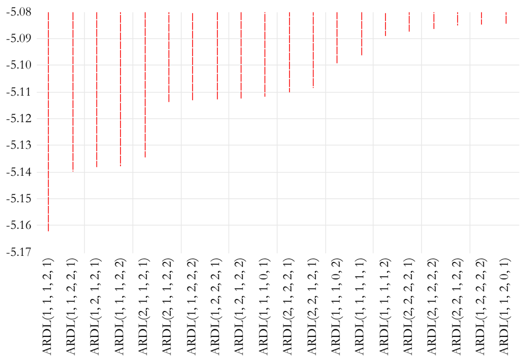
(i) Expenditure model (E1)



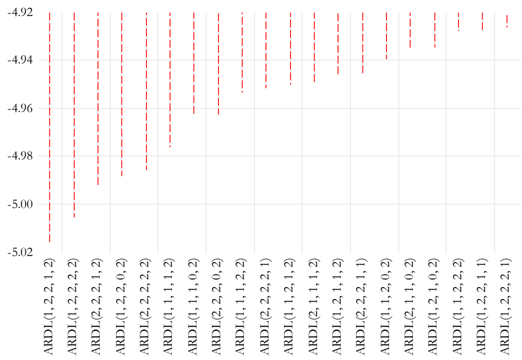
(iii) Revenue model (R1)



(ii) Expenditure model (E2)



(iv) Revenue model (R2)



Annex 4: Criteria values for ARDL cointegration bound tests

(i) Specifications with a trend variable

k	**10%**		**5%**		**2.5 %**		**1%**	
	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
2	3.38	4.02	3.88	4.61	4.37	5.16	4.99	5.85
3	2.97	3.74	3.38	4.23	3.80	4.68	4.30	5.23
4	2.68	3.53	3.05	3.97	3.40	4.36	3.81	4.92

Notes: k=number of non-fixed regressors; in our case the fixed regressors are the constant, the linear trend and the dummy variable; values from Pesaran et al. (2001).

(ii) Specifications without a trend variable

k	**10%**		**5%**		**2.5 %**		**1%**	
	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
4	2.45	3.52	2.86	4.01	3.25	4.49	3.74	5.06
5	2.26	3.35	2.62	3.79	2.96	4.18	3.41	4.68

Notes: k=number of non-fixed regressors; in our case the fixed regressors are the constant, the linear trend and the dummy variable; values from Pesaran et al. (2001).