MASTER

MONETARY AND FINANCIAL ECONOMICS

MASTER´S FINAL WORK

DISSERTATION

OPTIMAL LAG WHEN ANALYSING THE FISCAL SUSTAINABILITY THROUGH THE COINTEGRATION APPROACH

ANA SOFIA SILVA SANTOS

JULY - 2025

MASTER

MONETARY AND FINANCIAL ECONOMICS

MASTER´S FINAL WORK

DISSERTATION

OPTIMAL LAG WHEN ANALYSING THE FISCAL SUSTAINABILITY THROUGH THE COINTEGRATION APPROACH

ANA SOFIA SILVA SANTOS

SUPERVISION:

JOSÉ ALVES

JULY - 2025

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor, Professor José Alves, for his guidance, advice, and availability.

I am immensely grateful to my parents for all the opportunities they have given me in life, and without whom I could not have reached this stage.

Finally, to all my friends who have always been there for me and kept me motivated in the most difficult times.

ABSTRACT

The present Dissertation studies the fiscal sustainability in European Union countries by analyzing the relationship between government revenue and expenditure, and the selection of the optimal lag length. Using annual data from 1995 to 2024 for the 27 European Union countries, vector autoregressive (VAR) models and Granger causality tests are applied. The optimal lag for each country is selected based on standard information criteria (AIC, HQIC, and SBIC). The results show that, in many countries, revenue Granger-causes expenditure; causality in the opposite direction is less frequent. Regarding optimal lags, it varies considerable across country – in this study the range considered was one to ten. These findings highlight the importance of considering lag structure when examining fiscal sustainability.

JEL: C33; C52; E62; H63, H68.
Keywords: Fiscal sustainability; fiscal reaction function; cointegration; optimal lag.

RESUMO

A presente dissertação estuda a sustentabilidade orçamental nos países da União Europeia, analisando a relação entre as receitas e as despesas públicas e a seleção do comprimento ótimo de *lag*. Utilizando dados anuais de 1995 a 2024 para os 27 países da União Europeia, são aplicados modelos *vector autoregressive* (VAR) e testes de causalidade de Granger. O *lag* ótimo para cada país é selecionado com base em critérios de informação padrão (AIC, HQIC e SBIC). Os resultados mostram que, em muitos países, as receitas são a causa direta das despesas; a causalidade no sentido oposto é menos frequente. No que respeita aos desfasamentos ideais, estes variam consideravelmente de país para país – neste estudo, o intervalo considerado foi entre um e dez. Estes resultados sublinham a importância de considerar a estrutura de *lags* quando se examina a sustentabilidade orçamental.

JEL: C33; C52; E62; H63, H68.

Palavras-chave: Sustentabilidade orçamental; função de reação orçamental; cointegração; *lag* ótimo.

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

One aspect that takes an important place in the analysis of a country's economy is its fiscal sustainability. An economy is considered fiscally sustainable when its government's revenues and expenditures are close to balance.

There are various approaches to assessing the fiscal sustainability of an economy. The one that will serve as the basis for this study and on which most articles focus is the relationship between revenue and expenditure.

Studies that analyze the long-term relationship (cointegration) between both variables show that there is a sustainable relationship between revenue and expenditure – the coefficient that represents the return obtained for each euro the government spends is very close to 1, as desired. However, when we move on to a contemporary analysis of revenue as a function of public expenditure, since lags are not taken into account, the relationships are weak – this result is to be expected, since the government, by spending 1 euro today, will not receive a significant fraction of that euro in the same period since the economy is not supposed to react immediately. This is the problem with the time-varying perspective.

Thus, the aim of this research, rather than analyzing over the long term or year by year is to find the optimal lag of the relationship between public revenue and expenditure in the cointegration relationship, covering the sustainability factor – whether it takes the government more or less time to get its money back.

The determination of the optimal lag is a crucial step in time series analysis, as it defines the number of periods required for the effect of one variable to be transmitted to another. By identifying the optimal lag between government revenue and expenditure, it is possible to have a more accurate understanding of the speed at which fiscal adjustments occur; on the other hand, choosing and inappropriate lag length can lead to biased estimates or misleading conclusions regarding causality in fiscal sustainability. Therefore, the optimal lag in this study is found using information criterion such as the Akaike information criteria (AIC), the Schwarz-Bayes criterion (BIC), and the Hannan-Quinn information criterion (HQIC).

The analysis carried out in this thesis covers the data of the countries of the European Union between 1995 and 2024.

The paper is organized as follows. Chapter 2 provides the literature review on the theme. Chapter 3 contains the empirical approach where the methodology and data are explained. Chapter 4 presents the empirical results. Finally, chapter 5 concludes.

2. LITERATURE REVIEW

2.1. FISCAL SUSTAINABILITY

Approaches to measuring fiscal sustainability fall into two categories – forward looking and backward looking. The former use forecasts to assess whether an economy is on the path to sustainability and, as such, the downside is the possibility of including some forecasting errors. With regard to the latter, on which this study focus, the aim is to analyze historical data to determine whether policies on public debt and the primary balance have, in the past, aligned with the intertemporal budget constraint (IBC). For their part, backward-looking studies can make use of a fiscal reaction function – Bohn framework –, and the analysis of a cointegration relationship.

Afonso and Rault (2010) investigated fiscal sustainability in the EU-15 countries between 1970 and 2006 through panel data techniques. Their study employed stationarity and cointegration tests in order to evaluate the long-term sustainability of public finances. The conclusion was that, while fiscal sustainability is a key issue in certain countries, overall, fiscal policy remains sustainable within the EU-15, particularly during two time periods (1970-1991 and 1992-2006). The authors' findings also suggest that fiscal adjustments were generally effective at maintaining the fiscal balance during the period under analysis. Still within the European Union, Afonso and Coelho (2022) focused on analyzing the fiscal sustainability determinants and the primary balance responses to debt.

For the United States, Aldama and Creel (2019) assessed its fiscal sustainability using a Markov-switching fiscal rule model with annual data from 1940. This framework showed that fiscal policy alternates between sustainable and unsustainable periods – which cannot be observed using constant-parameter models. This research highlighted that delayed fiscal adjustments may ensure sustainability if the reaction is strong enough.

Debrun et al. (2019) examined the difficulties of assessing public debt sustainability, highlighting its forward-looking and strategic nature. The authors argued that simple indicators such as the debt-to-GDP ratio are not sufficient, as sustainability also depends on political willingness to repay, market perceptions, and the structure of the debt. It was also highlighted the evolution of IMF debt sustainability frameworks – they now consider uncertainty, probabilistic tools, and stress testing.

2.2. FISCAL REACTION FUNCTION

Bohn (1998) proposed an empirical test to assess the sustainability of public debt, introducing the idea that fiscal policy can be analysed through a fiscal reaction function. Under this framework, the fiscal policy is sustainable if there is a positive response of the primary balance in one year to the public debt-to-GDP ratio accumulated by the end of the previous year. This approach consists of observing whether the primary balance reacts positively to increases in lagged debt, i.e. whether the government increases the primary balance today compared to yesterday's debt.

Almost a decade later, and from the same author, Bohn (2007) questioned the generalized reliance on unit root and cointegration tests as necessary conditions for assessing fiscal sustainability. It is demonstrated that the intertemporal budget constraint (IBC) can still be satisfied even if debt, revenues, and expenditures are integrated to an arbitrarily high order and not cointegrated. Instead, the author defends a behavioral approach based on error-correction type fiscal reaction functions, in which the primary balance adjusts in response to past debt levels.

Afonso et al. (2021) estimated fiscal reaction functions to a sample of 28 countries of the European Union over the period from 1995 to 2021. In line with Afonso (2005) and Afonso and Coelho (2022), this study confirmed the existence of the Ricardian fiscal regime – particularly after the global financial crisis, where it was verified that improvements in the primary balance drive further reductions in the public debt-to-GDP ratio if the long-term real interest rate is higher than the economic growth rate. This analysis is in accordance with Bohn (1998), whose theoretical framework stands on the idea that fiscal sustainability is achieved through a positive adjustment of the primary balance in response to rising public debt. Still within the study of European countries, there are other papers that have applied the method of estimating fiscal reaction functions, such as Gali and Perotti (2003), Fincke and Greiner (2011), and Checherita-Westphal and Žďárek (2017).

In their study, Ghosh et al. (2013) defined fiscal space as the difference between current debt and a country-specific debt limit, which varies widely across countries. The authors developed a model that introduces "fiscal fatigue", showing that the governments' capacity to raise primary balances weakens as debt approaches 90-100% of GDP. This study's findings suggest that this capacity of raising the primary balance in response to rising debt levels cannot last indefinitely, similarly to what was found in Checherita-Westphal and Žďárek (2017).

In order to examine the impact of inflation on fiscal sustainability, Afonso et al. (2023b) applied a two-step approach, in which the first step consisted of estimating a time-varying fiscal response coefficient within a fiscal reaction function framework; in the second step the authors assessed how different measures of inflation affect the coefficient obtained. Their findings show significant volatility in fiscal sustainability across countries and over time, usually deteriorating during economic downturns due to countercyclical policies. It also suggests that while inflation can relieve fiscal pressures in the short run, its effect on the revenue-expenditure relationship depends on how governments adjust policies over time.

Afonso and Alves (2022) analysed the impact of government spending efficiency on fiscal sustainability in 35 OECD countries between 2007 and 2020. The authors used time-varying fiscal reaction functions in order to estimate the responsiveness of government revenues to expenditures and to link these coefficients to public sector efficiency scores obtained through DEA. Their findings demonstrate that efficiency improvements are associated with stronger fiscal sustainability, particularly in Euro Area countries. For OECD countries, the literature has also the studies of Saadaoui et al. (2022), Afonso et al. (2023a) and Afonso et al. (2024), that assessed the fiscal sustainability of this group through the estimation of fiscal reaction functions.

2.3. COINTEGRATION BETWEEN REVENUE AND EXPENDITURE

Alternatively, another way of assessing fiscal sustainability is to look at the cointegration between public revenues and expenditures.

The vast majority of the literature says that a long-run relationship exists between government revenues and expenditures (e.g. Magazzino et al. 2019 for the G-7 countries, Afonso 2005, Brady and Maggazino 2018 and Afonso and Coelho 2022 for the European Union countries).

Hakkio and Rush (1991) concluded that the existence of cointegration between government revenue and expenditure is a necessary condition for fiscal sustainability when both series are non-stationary. The author's findings suggest that if such a relationship exists, the government satisfies its present value budget constraint (PVBC), which ensures that deficits do not grow unsustainably over time. However, in the case of no integration between revenues and expenditures, it indicates potential fiscal imbalances in the long-term.

Another of the earliest empirical tests of fiscal sustainability using cointegration techniques was presented by Trehan and Walsh (1991). The authors analyzed the cointegration relationships between the United States government revenues and expenditures and concluded that there both variables are cointegrated – this would imply that fiscal policy is consistent with intertemporal budget constraints. Still within the United States, other studies that worked on the cointegration between revenues and expenditures are Quintos (1995) and Elmendorf and Mankiw (1999). The first argues that cointegration between revenues and expenditures is sufficient, however not necessary for a fiscally sustainable scenario. Using data with structural break tests, the author found that even when cointegration fails (especially after the early 1980s), the deficit process can still be sustainable under a weaker condition, as long as debt grows more slowly than the interest rate.

Afonso and Jalles (2017) examined fiscal sustainability in 11 Euro Area countries using quarterly panel data between 1991Q1 and 2013Q4. This study assessed the long-term relationship between the primary balance and government debt, and through the unit root and cointegration tests, the author's found results of sustainability in some countries, such as Germany, France, Belgium and the Netherlands, while for others there was no firm evidence. Time-varying coefficient models were also estimated in order to analyze the evolution of fiscal responsiveness, highlighting how fiscal sustainability can change over time.

From a wider cross-country perspective, Afonso et al. (2022) estimated a time-varying cointegration relationship between government revenues and expenditures as a proxy for fiscal sustainability. The authors used expanding-window regressions across 189 countries between 1980 and 2018 in order to derive a country-specific coefficient that would reflect the long-run alignment of fiscal variables – the closer this coefficient was to one, the more sustainable public finances would be. This study's findings suggest that the implementation of fiscal consolidations improves the stability of this long-run relationship, particularly in advanced economies.

2.4. OPTIMAL LAGS

The choice of the optimal number of lags is crucial when modeling cointegration relationships between public revenues and expenditures. The method employed in this thesis to reach the optimal lags was the consideration of traditional information criteria such as AIC, HQCI and SBIC (see Carrasco Gutiérrez et al. 2009 for an analysis on this topic).

An important issue that impacts on fiscal sustainability is precisely the timing of policy responses, hence the importance of estimating the optimal lags. This was noted in Auerbach (1994), where the author investigated how delays in fiscal adjustments, political incentives, and uncertainty can undermine the effectiveness of fiscal policy. Rather than modelling lag structures explicitly, this study highlights the risks of inaction and delayed responses to growing debt.

Favero and Giavazzi (2007) analyzed fiscal sustainability through the examination of the dynamic effects of fiscal shocks on the public debt path. They highlighted the need for richer dynamic specifications that incorporate lagged debt terms, showing that omitting these dynamics can lead to unstable and unrealistic debt paths. By including the lagged debt-to-GDP, the authors showed that proper lag specification is relevant for obtaining stable and realistic results. This conclusion aligns with other findings in literature, such as Afonso and Jalles (2015) that assessed fiscal sustainability in 18 OECD countries from 1970 to 2010 incorporating lagged fiscal variables into their models – although the focus of their work was not the selection of optimal lags per se, it reinforces the idea that ignoring the temporal dimension of fiscal responses may lead to misleading conclusions. Similarly, Égert (2010) also does not focus directly on the selection of lags, however it is an example of including dynamic fiscal responses into empirical models.

3. EMPIRICAL APPROACH

The present section aims to expose the methodology and data applied to study fiscal sustainability through the cointegration approach.

3.1. METHODOLOGY

Firstly, to find the optimal lag between the two variables studied, a Vector Autoregression (VAR) model was applied. Additionally, in order to assess if one variable can be useful to forecasting the other, the method used was the Granger causality test.

Vector Autoregression (VAR) model is a statistical model employed in order to analyze the relationship between multiple variables over time. In a VAR model, each variable is expressed as a linear function of its own past values and the past values of all other variables in the system. In the present study, this model was used to find the optimal lag length of government revenue and expenditure for each country and for the sample as a whole.

The Granger causality test is a statistical hypothesis test applied to determine if a time series can predict another. A time series X is considered to Granger-cause Y if past values of X contain statistically significant information about future values of Y. It is important to note that Granger causality does not imply true causality, but rather a temporal predictive ability. In this study, the Granger causality test was carried out in order to analyze the relationship between government revenue and expenditure.

3.2. DATA

The data this dissertation is based on was collected from the AMECO database. The countries targeted by the study are those of the European Union: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden. The period considered is from 1995 to 2024. This results in a balanced panel dataset covering a total of 810 observations – 30 years of data for each country.

The two variables under analysis are total government revenue and total government expenditure, and are described as follows.

Government revenue → the total income received by the government to finance the goods and services provided to citizens and businesses. It includes tax revenues, social contributions, and non-tax revenues (e.g. sales from public enterprises). In this study, government revenue is expressed as a percentage of GDP.

Government expenditure → the total government consumption, investment, and transfer payments. It includes current spending (e.g. wages, subsidies) and capital expenditure (e.g. infrastructure development). In this study, government expenditure is expressed as a percentage of GDP.

The first step of the empirical approach consisted of calculating the descriptive statistics for both variables: mean, minimum, maximum, and standard deviation for each country and for the EU as a whole, as it is shown in Table 1 and Table 2. By doing this exercise, it is possible to assess the main patterns and dispersion in the data and to identity potential irregularities.

Table 1. Descriptive Statistics – Total Government Revenue (% of GDP)

Country	Obs.	Mean	Min	Max	Std. Dev.
EU	810	42.7	21.8	57.9	6.6
Austria	30	49.7	48.3	51.6	0.8
Belgium	30	49.9	48.1	52.7	1.1
Bulgaria	30	36.6	26.3	43.1	3.6
Croatia	30	45.7	41.0	57.4	3.9
Cyprus	30	38.7	31.5	44.3	3.8
Czechia	30	39.6	37.0	42.1	1.3
Denmark	30	53.9	48.3	56.5	1.8
Estonia	30	38.4	34.9	43.4	1.9
Finland	30	52.9	50.9	56.3	1.3
France	30	52.0	50.3	54.3	1.3
Germany	30	45.4	43.5	47.5	1.2

Greece	30	44.3	36.8	50.5	4.7
Hungary	30	44.1	41.0	48.4	1.9
Ireland	30	31.4	21.8	38.4	5.3
Italy	30	45.7	43.0	48.0	1.5
Latvia	30	37.8	33.4	43.9	2.4
Lithuania	30	34.9	32.3	38.3	1.8
Luxembourg	30	43.4	41.3	47.9	1.6
Malta	30	36.1	32.0	40.0	2.3
Netherlands	30	43.9	42.6	46.2	0.9
Poland	30	40.4	37.7	46.4	0.8
Portugal	30	41.5	37.4	44.7	2.1
Romania	30	32.7	29.6	35.4	1.4
Slovakia	30	38.7	33.6	44.4	3.0
Slovenia	30	44.7	43.1	46.7	0.9
Spain	30	38.8	34.8	42.8	1.9
Sweden	30	52.3	48.5	57.9	2.7

Note: Author's calculations

Table 1 shows that the average government revenue as a percentage of GDP varies substantially across EU countries. Northern and Western European countries tend to have higher revenue ratios – for example, Denmark, Finland and Sweden exhibit average revenues above 50% of GDP – while countries like Ireland and Romania are significantly below the EU average of 42.7%.

Table 2. Descriptive Statistics - Total Government Expenditure (% of GDP)

	Obs.	Mean	Min	Max	Std. Dev.
EU	810	45.4	20.6	64.9	6.9
Austria	30	52.5	49.1	57.3	2.2
Belgium	30	52.5	48.6	58.5	2.4
Bulgaria	30	37.7	31.8	43.2	3.2
Croatia	30	49.2	44.3	57.9	3.9
Cyprus	30	41.2	32.5	54.6	4.8
Czechia	30	42.8	38.5	52.8	2.8
Denmark	30	52.9	44.9	58.7	3.5
Estonia	30	38.7	33.5	46.1	3.3
Finland	30	53.1	46.6	61.0	3.9
France	30	56.2	52.6	61.7	2.2
Germany	30	47.3	43.5	55.2	2.6
Greece	30	50.5	45.1	63.9	4.6
Hungary	30	49.2	45.8	55.0	1.9
Ireland	30	33.7	20.6	64.9	9.2

Italy	30	49.6	46.4	56.8	2.8
Latvia	30	40.2	35.4	46.9	3.6
Lithuania	30	37.4	33.4	50.0	3.9
Luxembourg	30	41.8	37.4	47.0	2.5
Malta	30	39.9	32.7	45.2	2.8
Netherlands	30	45.5	42.1	54.8	2.6
Poland	30	44.4	41.0	50.9	2.4
Portugal	30	45.7	42.3	51.9	3.0
Romania	30	36.8	33.2	43.5	2.7
Slovakia	30	43.3	35.9	53.2	4.5
Slovenia	30	48.0	43.5	57.7	3.1
Spain	30	42.9	38.3	51.4	3.7
Sweden	30	52.6	48.9	63.2	3.5

Note: Author's calculations

Table 2 shows that this wide variety of values across countries is also evident in government expenditure. France, Finland and Sweden have some of the highest average public spending levels (above 52% of GDP), whereas Ireland, Malta and Romania display the lowest values.

The second step was to test the stationarity of the two series, since time series models such as VAR and Granger causality require that the underlying variables are either stationary or made stationary through transformations. To assess the stationarity, the Augmented Dickey-Fuller (ADF) test was applied to each variable, for each country, using three specifications: level first differences, and first differences with drift – if the null hypothesis of a unit root was not rejected at levels, the first difference was tested; in cases in which the first difference test remained inconclusive, the test was repeated with a drift term to capture potential trends. The results are reflected in Table 3 and Table 4.

Table 3. Dickey-Fuller Test – Total Government Revenue (% of GDP)

	Level	Obs.	First Differences	Obs.	First Diff w/ drift	Obs.
Austria	-2.096	28	-4.103**	27		
Belgium	-2.091	28	-3.476*	27		
Bulgaria	-4.078**	28				
Croatia	-1.641	28	-5.261***	27		
Cyprus	-2.148	28	-3.680**	27		
Czechia	-3.288*	28				
Denmark	-3.116	28	-4.230**	27		

Estonia	-2.747	28	-4.968***	27		
Finland	-2.464	28	-3.480*	27		
France	-1.604	28	-3.044**	27	-3.105***	27
Germany	-1.865	28	-5.470***	27		
Greece	-2,289	28	-3.540*	27		
Hungary	-2.075	28	-4.121**	27		
Ireland	-1.840	28	-2.205	27	-2.294**	27
Italy	-2.131	28	-3.686**	27		
Latvia	-2.844	28	-3.515*	27		
Lithuania	-2.729	28	4.939***	27		
Luxembourg	- 0.849	28	-3.829**	27		
Malta	-1.440	28	-2.596	27	-2.596***	27
Netherlands	-3.316*	28				
Poland	3.728**	28				
Portugal	-2.628	28	-4,008**	27		
Romania	-4.784***	28				
Slovakia	-1.486	28	-3.330*	27		
Slovenia	-2.669	28	-5.027***	27		
Spain	-2.544	28	-4.193**	27		
Sweden	-2.702	28	-3.609**	27		

Note: * indicates the level of significance of 10%, ** a level of 5% and *** a level of 1%. Obs. are the observations for each regression.

Table 3 presents the results of the Dickey-Fuller test for government revenue. It is possible to observe that most countries are not stationary at levels, but become stationary after first-differencing – this is an expected result in macroeconomic time series. Some countries (e.g. Bulgaria, Romania and the Netherlands) already display stationarity in levels. In only three cases, it was necessary to include a drift term in order to achieve stationarity.

Table 4. Dickey-Fuller Test – Total Government Expenditure (% of GDP)

	Level	Obs.	First Differences	Obs.	First Diff w/ drift	Obs.
Austria	-3.705**	28				
Belgium	-2.932	28	-4.560***	27		
Bulgaria	-3.581*	28				
Croatia	-2.735	28	-5.292***	27		
Cyprus	-1.664	28	-4.230**	27		
Czechia	-3.084	28	-4.653***	27		
Denmark	-2.072	28	-3.964**	27		
Estonia	-3.357*	28				
Finland	-3.435*	28				
France	-3.211	28	-4.678***	27	-	

Germany	-2.118	28	-5.187***	27		
Greece	-2.209	28	-3.973**	27		
Hungary	-3.696**	28				
Ireland	-1.689	28	-3.426*	27		
Italy	-3.536*	28				
Latvia	-3.328*	28				
Lithuania	-2.998	28	-5.988***	27		
Luxembourg	-3.131	28	-4.494***	27		
Malta	-2.513	28	-3.676**	27		
Netherlands	-2.907	28	-4.222**	27		
Poland	-2.373	28	-4.957***	27		
Portugal	-2.148	28	-5.555***	27		
Romania	-2.119	28	-3.058	27	-3.046***	27
Slovakia	-1.692	28	-4.748***	27		
Slovenia	-2.641	28	-4.164**	27		
Spain	-2.868	28	-3.714**	27		
Sweden	-3.551*	28				

Note: * indicates the level of significance of 10%, ** a level of 5% and *** a level of 1%. Obs. are the observations for each regression.

Table 4 shows the stationarity tests for government expenditure and, similarly to revenue, most countries required first-differencing to achieve stationarity; only a few cases, such as Hungary and Austria, already exhibited stationarity at level. Overall, the rest results supported the use of a VAR framework, since both variables were integrated of the same order (mostly I(1)).

Once the stationarity properties of both variables were confirmed, the next phase was to determine the optimal lag length for each country – a crucial part of this study and an important decision in time-series models, since the number of lags may influence the validity of the causality tests: short lag structures may omit relevant information, while excessively long ones can increase estimation errors.

To determine the optimal lag, a series of VAR models was estimated for each country using lag lengths from 1 to 10. The optimal lag was then selected using the three common information criteria: the Akaike Information Criterion (AIC), the Hannan-Quinn Criterion (HQIC), and the Schwarz Bayesian Information Criterion (SBIC). These criteria provide a balance between model fit and parsimony by penalizing the inclusion of excessive parameters. The AIC tends to favor more complex models, often suggesting longer lag structures, while the SBIC is more conservative and usually selects shorter lags due to its stronger penalty term. The HQIC lies between the two, offering a compromise between model accuracy and simplicity. The comparison of the results using these three criteria helps to ensure robustness in the selection of the lag length, capturing the most appropriate dynamic structure of the relationship between revenue and expenditure.

Table 7, presented in the Appendix, reports the AIC, HQIC, and SBIC values of the chosen lags for each country. These values correspond to the optimal lag selected for each case, rather than the full set of results for all ten lag lengths tested, in order to ensure a more concise presentation. In most cases, the optimal lag selected by the three criteria was consistent, indicating strong convergence in the model selection process; however, in other cases, it was necessary to contrast the choice with the p-values (statistical significance), if it would result in more robust results in the Granger causality tests.

Finally, the last step consisted of applying the Granger causality test in order to assess the relationship between the two variables. The objective was to find causality in both directions: to test whether revenue Granger-causes expenditure, or if expenditure Granger-causes revenue. This test aimed to analyze if past values of one variable contain information that helps to predict the other. The results are presented in the next chapter.

4. EMPIRICAL ANALYSIS

This chapter presents and analyses the empirical results obtained from the Granger causality tests between government revenue and government expenditure for the 27 EU countries. As previously indicated, the analysis was conducted using the optimal lag length selection via information criteria (AIC, HQIC and SBIC) and applied to each country using annual data from 1995 and 2024. The causality was tested in both directions: if revenues cause expenditures, and if expenditures cause revenues.

Table 5. Granger-causality Test – Revenues cause Expenditures

	Optimal lag	Obs.	rev_gdp causes exp_gdp p-value
Austria	4	26	0.047**
Belgium	3	27	0.008***
Bulgaria	5	25	0.004***
Croatia	10	20	0.071*
Cyprus	7	23	0.210
Czechia	7	23	0.000***
Denmark	1	29	0.023**
Estonia	10	20	0.027**
Finland	6	24	0.936
France	2	28	0.001***
Germany	2	28	0.066*
Greece	3	27	0.024**
Hungary	5	25	0.063*
Ireland	5	25	0.026**
Italy	8	22	0.036**
Latvia	3	27	0.053*
Lithuania	5	25	0.020**
Luxembourg	1	29	0.851
Malta	5	25	0.000***
Netherlands	5	25	0.003***
Poland	6	24	0.000***
Portugal	3	27	0.019***
Romania	8	22	0.117
Slovakia	7	23	0.006***
Slovenia	4	26	0.010**

Spain	8	22	0.000***
Sweden	4	26	0.000***

Note: * indicates the level of significance of 10%, ** a level of 5% and *** a level of 1%. Obs. are the observations for each regression.

Table 5 shows the results on testing if revenue can Granger-cause expenditure. In several countries, a statistically significant causal relationship was found at the 1% or 5% level. Sweden (lag 4, p = 0.000) and Netherlands (lag 5, p = 0.003) are examples of highly significance, meaning that the past values of revenue provide useful information in predicting expenditure values.

However, in four countries (Cypryus, Finland, Luxembourg and Romania) there was no evidence of Granger causality in this direction. Spain and Czechia, despite having a very low p-value (0.000), they have high lag length (lag 8 and 7, correspondently) which may raise questions about overfitting – this should be considered when interpreting the strength of the result.

Overall, more than half of the countries targeted display some degree of causality from revenue to expenditure.

Table 6. Granger-causality Test – Expenditures cause Revenues

	Optimal lag	**Obs.**	**exp_gdp causes rev_gdp p-value**
Austria	4	26	0,003***
Belgium	3	27	0,353
Bulgaria	5	25	0,031**
Croatia	10	20	0,018**
Cyprus	7	23	0,084*
Czechia	7	23	0,252
Denmark	1	29	0,932
Estonia	10	20	0,042**
Finland	6	24	0,000***
France	2	28	0,385
Germany	2	28	0,001
Greece	3	27	0,893
Hungary	5	25	0,045**
Ireland	5	25	0,002***

Italy	8	22	0,001***
Latvia	3	27	0,101
Lithuania	5	25	0,966
Luxembourg	1	29	0,034**
Malta	5	25	0,001***
Netherlands	5	25	0,090*
Poland	6	24	0,003***
Portugal	3	27	0,645
Romania	8	22	0,002***
Slovakia	7	23	0,001***
Slovenia	4	26	0,011**
Spain	8	22	0,013**
Sweden	4	26	0,025**

Note: * indicates the level of significance of 10%, ** a level of 5% and *** a level of 1%. Obs. are the observations for each regression.

Doing the exercise to the reverse relationship, Table 6 shows the results of testing if expenditure may Granger-cause revenue. Fewer countries exhibit significant causal links in this direction, though it still is the majority – out of the 27 analysed countries, only 8 do not display significance in this relationship. In Estonia and Greece, causality was found at the 5% level, and in Austria and Finland, ate 1%, suggesting that expenditure decisions may influence future revenue – possibly through automatic stabilizers or fiscal policies.

Another relevant point in this study concerns the optimal lag used for each country. While the majority of models followed the lag selected by at least two of the three information criteria, in some cases the second or even third-best lag was preferred. This adjustment is justified by the performance of the models employed, particularly when a lower-ranked lag had more statistically meaningful results. Having this flexibility allowed a balance between statistical robustness and theoretical consistency.

In the second column of Table 5 and Table 6, it is possible to observe the chosen lags for each country. As expected, the optimal lags vary substantially, ranging from lag 1 to 10, and this reflects the heterogeneity of fiscal dynamics in different national contexts. In general, countries with more stable or institutionalized fiscal frameworks tended to haver shorter optimal lags (e.g. Austria, Belgium), while others that present a more volatile fiscal structure, need longer lag structures (e.g. Croatia, Czechia).

Regarding the countries in which no Granger causality was found in either direction, this absence may indicate that both variables, government revenue and government expenditure, evolve independently. This can be due to rigid institutional conditions, European Union fiscal constraints, or political factors that weaken the response between the two variables.

Concluding, the evidence shows that Granger causality between revenue and expenditure is present in a considerable number of EU countries, although the strength and direction of the relationship may vary. The most common case is the one in which revenue tends to Granger-cause expenditure, which the perspective of a reactive fiscal stance where government adjust their spending in response to fiscal capacity. On the other hand, expenditure-driven causality is less frequent and this suggests limited evidence for an expenditure-led revenue strategy.

The next and final chapter summarizes the findings of this study.

5. CONCLUSIONS AND POLICY IMPLICATIONS

The objective of this dissertation was to study the fiscal sustainability in European Union countries by analyzing the relationship between government revenue and expenditure, with a particular focus on the optimal lag length in the context of Granger causality analysis. Using annual data from 1995 to 2024 for the 27 UE Member States, it was employed vector autoregressive (VAR) models and selected lag structures based on standard information criteria.

This study's results reveal a heterogeneous pattern of fiscal interactions across countries. In a significant part, government revenue Granger-causes expenditure, suggesting a reactive fiscal stance in which expenditure is adjusted in response to revenue developments. In contrast, evidence of causality in the opposite direction – from expenditure to revenue – is considerably weaker and less frequent. These findings are in line with the literature, which often highlights the prevalence of revenue-led fiscal adjustments in rules-based fiscal environments such as the EU.

A key contribution of this study concerns the emphasis placed on the optimal lag selection. The analysis demonstrated that lag lengths can vary widely across countries, ranging from one to ten years (the interval of lags considered). This variation highlights the importance of adapting empirical specifications to each country, as the choice of lag can significantly affect the interpretation of causality between variables. In some cases, second or third-best lags were selected to ensure better model fit and more robust statistical results.

It is important to note that this study has its limitations. Using annual data may be sufficient for long-term analysis, however it may hide short-run dynamics or fiscal adjustments covered in the same year. Additionally, this analysis focused on the relationship between two variables and did not incorporate other macroeconomic variables that might influence this interaction.

Overall, this Dissertation had the objective of understanding the fiscal dynamics in the European Union by focusing on the direction and timing of adjustments between government revenue and expenditure.

REFERENCES

1. Afonso, A. (2005). Fiscal Sustainability: The Unpleasant European Case. *FinanzArchiv / Public Finance Analysis*, *61*(1), 19–44. https://doi.org/10.1628/0015221053722532

2. Afonso, A., Alves, J., & Coelho, J. C. (2023a). Determinants of the degree of fiscal sustainability. *CESifo Working Paper*, No. 10225. https://doi.org/10.2139/ssrn.4338242

3. Afonso, A. & Alves, J. (2022). Does government spending efficiency improve fiscal sustainability? *CESifo Working Paper*, No. 9738. https://doi.org/10.2139/ssrn.4103951

4. Afonso, A., Alves, J., & Jalles, J. T. (2021). Fiscal Reaction Functions in the EU: A Tale of “r–g” and Whether it Matters. *Révue française d'économie*, XXXVI, 65-101. https://doi.org/10.3917/rfe.213.0065

5. Afonso, A., Alves, J., & Jalles, J. T. (2022). To consolidate or not to consolidate? A multi-step analysis to assess needed fiscal sustainability. *CESifo Working Paper*, No. 9618. http://dx.doi.org/10.2139/ssrn.4054123

6. Afonso, A., Alves, J., Matvejevs, O. & Tkacevs, O. (2023b). Fiscal Sustainability and the Role of Inflation. *CESifo Working Paper*, No. 10843. https://doi.org/10.2139/ssrn.4676761

7. Afonso, A., Alves, J., Matvejevs, O., & Tkačevs, O. (2024). Impact of sovereign debt maturity on fiscal sustainability. *CESifo Working Paper*, No. 11499. https://doi.org/10.2139/ssrn.5051009

8. Afonso, A. & Coelho, J. C. (2022). Fiscal sustainability, fiscal reactions, pitfalls and determinants. *CESifo Working Paper*, No. 9635. https://doi.org/10.2139/ssrn.4063444

9. Afonso, A. & Jalles, J. T. (2015). Fiscal sustainability: A panel assessment for advanced economies. *Applied Economics Letters, 22*(11), 925–929. https://doi.org/10.1080/13504851.2014.987913

10. Afonso, A. & Jalles, J. T. (2017). Euro Area Time-Varying Fiscal Sustainability. *International Journal of Finance and Economics*, *22*(3), 244-254. https://doi.org/10.1002/ijfe.1582

11. Afonso. A. & Rault, C. (2010). What do we really know about fiscal sustainability in the EU? A panel data diagnostic. *Review of World Economics, 145*(4), 731‑755. https://doi.org/10.2139/ssrn.1996486

12. Aldama, P. & Creel, J. (2019). Fiscal policy in the US: Sustainable after all? *Economic Modelling, 81*(C), 471-479. https://doi.org/10.1016/j.econmod.2018.03.017

13. Auerbach, A. J. (1994). The U.S. Fiscal Problem: Where We Are, How We Got Here, and Where We're Going. *NBER Working Paper*, No. 3656. https://doi.org/10.2307/3585081

14. Bohn, H. (1998). The Behavior of U.S. Public Debt and Deficits. *The Quarterly Journal of Economics 113*(3), 949‑963. https://doi.org/10.1162/003355398555793

15. Bohn, H. (2007). Are stationarity and cointegration restrictions really necessary for the intertemporal budget constraint? *Journal of Monetary Economics, 54*(7), 1837-1847. https://doi.org/10.1016/j.jmoneco.2006.12.012

16. Brady, G. L. & Magazzino, C. (2018). Fiscal Sustainability in the EU. *Atlantic Economic Journal*, *46*, 297-311. https://doi.org/10.1007/s11293-018-9588-4

17. Carrasco Gutiérrez, C. E., Souza, R. C. & Guillén, O. T. de C. (2009). Selection of Optimal Lag Length in Cointegrated VAR Models with Weak Form of Common Cyclical Features. *Brazilian Review of Econometrics*, *29*(1), 1-14. https://doi.org/10.12660/ bre.v29n12009.2696

18. Checherita-Westphal, C. & Žďárek, V. (2017). Fiscal reaction function and fiscal fatigue: evidence for the euro area. *Working Paper Series 2036*, *European Central Bank*. https://doi.org/10.2866/827938

19. Debrun, X., Ostry, J. D., Willems, T. & Wyplosz, C. (2019). Public Debt Sustainability. *CEPR Discussion Papers*, DP14010. Centre for Economic Policy Research. https://doi.org/10.1093/oso/9780198850823.003.0005

20. Égert, B. (2010). Fiscal Policy Reaction to the Cycle in the OECD: Pro- or Counter-Cyclical? *OECD Economics Department Working Papers*, 763, OECD Publishing. https://dx.doi.org/10.1787/5kmft7pthb27-en

21. Elmendorf, D. W., & Mankiw, N. G. (1999). Government debt. In J. B. Taylor & M. Woodford (Eds.), *Handbook of Macroeconomics*, *1*, 1615-1669. Elsevier. https://doi.org/10.1016/s1574-0048(99)10038-7

22. Favero, C. A., & Giavazzi, F. (2007). Debt and the effects of fiscal policy. *National Bureau of Economic Research, NBER Working Paper,* No. 12822. http://dx.doi.org/10.2139/ssrn.1002383

23. Fincke, B. & Greiner, A. (2011). Debt Sustainability in Selected Euro Area Countries: Empirical Evidence Estimating Time-Varying Parameters. *Studies in Nonlinear Dynamics & Econometrics*, *15*(3), 1-21. https://doi.org/10.2202/1558-3708.1777

24. Gali, J., & Perotti, R. (2003). Fiscal policy and monetary integration in Europe. *Economic Policy*, *18*(37), 533–572. https://doi.org/10.3386/w9773

25. Ghosh, A. R., Kim, J. I., & Mendoza, E. G., Ostry, J. D., & Qureshi, M. S. (2013). Fiscal fatigue, fiscal space and debt sustainability in advanced economies. *Economic Journal*, *123*(566), F4–F30. https://doi.org/10.1111/ecoj.12010

26. Hakkio, C. S. & Rush, M. (1991). Is the Budget Deficit "Too Large"?. *Economic Inquiry, 29*(3), 429-445. https://doi.org/10.1111/j.1465-7295.1991.tb00837.x

27. Magazzino, C., Brady, G. L., & Forte, F. (2019). A panel data analysis of the fiscal sustainability of G-7 countries. *The Journal of Economic Asymmetries 20*(C), e00127. https://doi.org/10.1016/j.jeca.2019.e00127

28. Quintos, C. E. (1995). Sustainability of the deficit process with structural shifts. *Journal of Business & Economic Statistics, 13*(4), 409–417. https://doi.org/10.1080/07350015.1995.10524615

29. Saadaoui, J., Lau, C. K. M., & Cai, Y. (2022). Testing fiscal sustainability in OECD countries: new evidence from the past centuries. *Applied Economics Letters, 31*(7), 676–682. https://doi.org/10.1080/13504851.2022.2142502

30. Trehan, B., & Walsh, C. E. (1991). Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits. *Journal of Money, Credit and Banking*, *23*(2), 206–223. https://doi.org/10.2307/1992777

APPENDIX

Table 7. Vector Autoregression (VAR) – lag selection criterium

	Lags	AIC	SBIC	HQIC
Austria	4	6.589446	6.879776	6.673051
Belgium	3	7.453612	7.741575	7.539238
Bulgaria	5	8.932688	9.225218	9.013824
Croatia	10	8.089905	8.388625	8.148218
Cyprus	7	9.679839	9.976055	9.754337
Czechia	7	6.874975	7.171191	6.949473
Denmark	1	7.659098	7.941987	7.747695
Estonia	10	8.136406	8.435126	8.194719
Finland	6	7.611822	7.906336	7.689957
France	2	6.708766	6.994238	6.796038
Germany	2	6.637688	6.92316	6.72496
Greece	3	10.39273	10.68069	10.47836
Hungary	5	8.084519	8.377049	8.165654
Ireland	5	12.46235	12.75488	12.54349
Italy	8	7.848128	8.145685	7.918223
Latvia	3	9.291336	9.5793	9.376963
Lithuania	5	8.095491	8.388021	8.176626
Luxembourg	1	7.118307	7.401196	7.206904
Malta	5	8.241805	8.534335	8.322941
Netherlands	5	5.986896	6.279426	6.068032
Poland	6	7.17495	7.469463	7.253084
Portugal	3	8.153547	8.239173	8.44151
Romania	8	7.845797	8.143354	7.915892
Slovakia	7	8.879193	9.175409	8.953691
Slovenia	4	7.064538	7.354868	7.148143
Spain	8	9.183873	9.48143	9.253969
Sweden	4	6.270277	6.560607	6.353882

Note: Author’s calculations