The sustainability of fiscal policies: a study of the European Union

This item was submitted to Loughborough University’s Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University

Metadata Record: [https://dspace.lboro.ac.uk/2134/10548](https://dspace.lboro.ac.uk/2134/10548)

Publisher: © Carlos Vieira

Please cite the published version.
This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (https://dspace.lboro.ac.uk/) under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
The Sustainability of Fiscal Policies:
A Study of the European Union

by

Carlos Manuel Rodrigues Vieira

A Doctoral Thesis
Submitted in partial fulfilment
of the requirements for the award of
Doctor of Philosophy

Department of Economics
Loughborough University
September 1999

© Carlos Vieira
ABSTRACT

The concern with persistent high government deficits and debts has been one of the most controversial and discussed issues among academics and policymakers during the last two decades of the twentieth century. Despite recent efforts towards fiscal consolidation in most developed countries, expensive welfare programs and unfunded social security systems can exert a considerable strain on public finances over the next generations.

The main objective of this thesis is to investigate whether current fiscal policies are sustainable, that is, able to guarantee the government's solvency, and what are the consequences of unsustainability on monetization, inflation and interest rates. The first question is tested by examining the long-run univariate and multivariate stochastic properties of the fiscal variables, as implied by the intertemporal budget constraint. The second question is assessed within a vector autoregressive framework, which allows the consideration of feedback mechanisms often neglected in the literature. More specifically, the econometric methodology employed throughout the study comprises recent developments in cointegration analysis, panel data techniques, bounds-ARDL procedure, and Granger non-causality.

The empirical analysis is focused on a comparative study of six core members of the European Union, during the post-war period: Belgium, France, Germany, Italy, Netherlands and United Kingdom. The evidence suggests that only Germany and the Netherlands have been following a sustainable fiscal path, although the latter remains vulnerable to the consequences of an ever-increasing stock of debt. However, unsustainable fiscal policies do not seem to have imposed an excessive burden on monetary policies, as predicted by the conventional economic theory. Apart from Italy, there is no empirical evidence that high deficits necessarily imply monetary financing, growing inflation and rising interest rates.

KEYWORDS: fiscal sustainability, intertemporal budget constraint, government debt, deficits, monetization, interest rates, EMU.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Professor Christopher J. Green and Dr. Mark J. Holmes, a constant source of assistance, guidance and motivation throughout the course of the research. Their intellectual input and academic rigour have been invaluable. The faults of the work are, however, my sole responsibility.

My gratitude is extended to all the academic and support staff in the Department of Economics, and also to my fellow students, for providing an excellent working environment. I am particularly grateful to the Head of Department, Professor Tom Weyman-Jones, for his warm welcome and unswerving encouragement, to my Director of Research, Professor Terence C. Mills, for his support and econometric advice, and also to Su Spencer, for her generous help.

I have also benefited from the discussion, help and incentive of many colleagues from the Universidade de Évora. Special thanks to António Caleiro, Pedro Damião, Claudia Padrão, Fernanda Peixe, Conceição Rego, Agostinho Rosa and also to José Rui Figueira from the Universidade de Coimbra.

In the process of data collection, many debts are usually incurred. I am under particular obligations to Ms Bethany Blowers from the Bank of England, for elucidating the role of the social security funds in the UK's central government budget, to L. Elfferich from the De Nederlandsche Bank and to Mr. Claude Modart from the National Bank of Belgium, for government's financial data, and to Dr Kastrop from the Bundesministerium der Finanzen, for explaining the German budget during the first post-war years.

I gratefully acknowledge the financial support from the Comissão Invotan and the Universidade de Évora.

On a more personal note, I want to thank all my friends and family, who have always been supportive despite the many miles of land and water getting in the way. To António Bicas, who has shown genuine friendship. To Hélder, Antónia and Zé Viegas, my second family. To Isabel, who shared this parallel venture and everything else in my life. To my brother Pedro and to my parents, Manuel Filipe and Maria Alice, for everything.
CONTENTS

List of tables vii
List of acronyms and abbreviations ix
List of notation xi

1. Introduction 1

2. The IBC and Alternative Sustainability Indicators 8
  2.1. Introduction 8
  2.2. The intertemporal budget constraint 9
  2.3. Sustainability with a collateral constraint 18
    2.3.1. Stabilisation of the current debt-GDP ratio 19
    2.3.2. Sustainability in the Maastricht Treaty and the Stability Pact 21
    2.3.3. Blanchard’s sustainability indicators 24
  2.4. Sustainability with a transversality condition 27
    2.4.1. The assumption of a constant interest rate 29
      2.4.1.1. The first tests: stationarity of the net-of-interest deficit 30
      2.4.1.2. Stationarity of the total deficit 32
      2.4.1.3. The Hansen, Roberds and Sargent test 37
    2.4.2. A variable interest rate 41
      2.4.2.1. Stationarity of the discounted debt (Wilcox’s test) 42
      2.4.2.2. Stationarity of the first-differenced debt 44
    2.4.3. The sustainability tests in a stochastic economy 48
  2.5. Summary 56

3. Methodology and Data 61
  3.1. Methodology adopted 61
    3.1.1. A sequential testing strategy 61
    3.1.2. The econometric tests 64
  3.2. Data description 74
    3.2.1. Main characteristics of the data sets 75
    3.2.2. Variables’ description, sources and adjustments 80
    3.2.3. The evolution of the fiscal stance in an historical context 88
6.4. Data and methodology

6.4.1. Choice of variables and other data issues

6.4.2. Methodology adopted to test the relationship

6.5. The relation deficits - interest rates

6.6. An analysis of sovereign risk

6.7. General government fiscal variables and the interest rate

6.8. The interest rate differentials: fiscal discipline through the financial markets

6.9. European-wide effects

6.10. Summary of the main results

7. Conclusion

Appendix

References
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The dynamics of the debt-GDP ratio</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary of the results for the European Union</td>
<td>60</td>
</tr>
<tr>
<td>3.1a</td>
<td>Definition, adjustments and sources of the variables - government finance</td>
<td>82</td>
</tr>
<tr>
<td>3.1b</td>
<td>Definition, adjustments and sources of the variables - national accounts</td>
<td>86</td>
</tr>
<tr>
<td>3.1c</td>
<td>Definition, adjustments and sources of the variables - interest rates</td>
<td>87</td>
</tr>
<tr>
<td>3.1d</td>
<td>Definition, adjustments and sources of the variables - other variables</td>
<td>88</td>
</tr>
<tr>
<td>3.2</td>
<td>The average growth-corrected real interest rate ($\rho$)</td>
<td>90</td>
</tr>
<tr>
<td>4.1</td>
<td>Unit root tests on the series of total expenditures and revenues ($t_g, t_t$)</td>
<td>96</td>
</tr>
<tr>
<td>4.2</td>
<td>Unit root tests on the series of total surplus ($t_s$)</td>
<td>98</td>
</tr>
<tr>
<td>4.3</td>
<td>Unit root tests allowing a structural break in the series</td>
<td>103</td>
</tr>
<tr>
<td>4.4</td>
<td>Unit root tests allowing two mean shifts</td>
<td>107</td>
</tr>
<tr>
<td>4.5</td>
<td>Engle-Granger cointegration tests</td>
<td>110</td>
</tr>
<tr>
<td>4.6</td>
<td>Johansen cointegration test</td>
<td>112</td>
</tr>
<tr>
<td>4.7</td>
<td>Tests on the parameters of the error correction terms</td>
<td>113</td>
</tr>
<tr>
<td>4.8</td>
<td>Tests of non-causality between revenues and expenditures</td>
<td>116</td>
</tr>
<tr>
<td>4.9</td>
<td>Average deficit of the general and central govts. in the period 1950-96</td>
<td>120</td>
</tr>
<tr>
<td>4.10</td>
<td>Unit root tests on the revenues and expenditures series</td>
<td>121</td>
</tr>
<tr>
<td>4.11</td>
<td>Unit root tests on the total surplus series</td>
<td>122</td>
</tr>
<tr>
<td>4.12</td>
<td>Unit root tests allowing a structural break in the series</td>
<td>122</td>
</tr>
<tr>
<td>4.13</td>
<td>Engle-Granger cointegration tests</td>
<td>123</td>
</tr>
<tr>
<td>4.14</td>
<td>Johansen's maximum eigenvalue cointegration test ($\lambda_{MAX}$ test)</td>
<td>125</td>
</tr>
<tr>
<td>4.15</td>
<td>Tests on the parameters of the error correction terms</td>
<td>126</td>
</tr>
<tr>
<td>4.16</td>
<td>Tests of non-causality between revenues and expenditures</td>
<td>127</td>
</tr>
<tr>
<td>4.17</td>
<td>ADF$_{ST}$ unit root tests on total revenues and expenditures, seasonally adjust.</td>
<td>129</td>
</tr>
<tr>
<td>4.18</td>
<td>Johansen cointegration tests, seasonally adjusted data</td>
<td>130</td>
</tr>
<tr>
<td>4.19</td>
<td>HEGY unit root tests on the fiscal variables</td>
<td>134</td>
</tr>
<tr>
<td>4.20</td>
<td>Long-run and seasonal unit root tests with a structural break</td>
<td>136</td>
</tr>
<tr>
<td>4.21</td>
<td>Engle-Granger and Johansen cointegration tests (filtered series)</td>
<td>138</td>
</tr>
<tr>
<td>4.22</td>
<td>Levin-Lin panel data unit root tests</td>
<td>141</td>
</tr>
<tr>
<td>4.23</td>
<td>Panel data unit root tests, summary of all possible combinations</td>
<td>142</td>
</tr>
<tr>
<td>4.24</td>
<td>Im, Pesaran and Shin's panel data unit root tests</td>
<td>145</td>
</tr>
</tbody>
</table>
4.25 Unit root and cointegration tests on the aggregate fiscal variables 147
5.1 Average value of seigniorage during the sample period (% GDP) 162
5.2 Sustainability tests on the non-monetization constraint 164
5.3 ADF$_{ST}$ unit root tests on the series of seigniorage and inflation 166
5.4 Bounds-ARDL test procedure applied to the relation $t_{s_{t}}, \sigma_{t}^{CF}$ 170
5.5 Bounds-ARDL test procedure applied to the relation $t_{s_{t}}, \sigma_{t}^{OC}$ 172
5.6 Johansen cointegration tests on seigniorage $(\sigma_{t}^{OC})$ and total surplus $(t_{s_{t}})$ 173
5.7 LA-VAR non-causality tests: seigniorage $(\sigma_{t}^{CF}, \sigma_{t}^{OC})$ and total surplus $(t_{s_{t}})$ 174
5.8 Bounds-ARDL test procedure applied to the relation $t_{s_{t}}, \pi_{t}^{CPI}$ 176
5.9 Johansen cointegration tests - total surplus $(t_{s_{t}})$ and inflation $(\pi_{t}^{CPI})$ 177
5.10 LA-VAR non-causality tests - inflation $(\pi_{t}^{CPI})$ and total surplus $(t_{s_{t}})$ 178
5.11 ADF$_{ST}$ unit root tests on the (log of) tax rate and base money to GDP ratio 185
5.12 Johansen cointegration tests - inflation $(\pi_{t}^{CPI})$, tax rate $ln t_{t}$, base money $ln m_{t}$ 186
6.1 Unit root tests on the series of debt and interest rates 206
6.2 Possible outcomes of the tests 209
6.3 Central government total surplus and the interest rates 212
6.4 Central government debt and the interest rates 216
6.5 General government total surplus and nominal interest rate 219
6.6 Central government total surplus and nominal interest rate differential 222
6.7 Central government aggregated fiscal variables and interest rates 226
6.8 The relation deficits-interest rates, summary results 227
6.9 The relation debt-interest rates, summary results 229
List of Acronyms

ADF  Augmented Dickey-Fuller
ADF_{MAX}  ADF, lag length chosen according to the Leybourne (1995) procedure
ADF_{SBC}  ADF, lag length chosen according to the SBC
ADF_{ST}  ADF, lag length chosen according to a sequential testing procedure
AIC  Akaike Information Criterion
AO  additive outlier
ARDL  autoregressive, distributed lag
ARIMA  autoregressive integrated moving average
ARMA  autoregressive moving average
B  Belgium
EEC  European Economic Community
EMS  European Monetary System
EMU  Economic and Monetary Union
ERM  Exchange Rate Mechanism
EU  European Union
F  France
G  Germany
HEGY  Hylleberg, Engle, Granger and Yoo (1990)
I  Italy
IBC  intertemporal budget constraint
IFS  International Financial Statistics
IFSY  International Financial Statistics Yearbook
IMF  International Monetary Fund
IO  innovative outlier
IPS  Im, Pesaran and Shin (1997)
LA-VAR  lag-augmented vector autoregression
LL  Levin and Lin (1992 and 1993)
N  Netherlands
OECD  Organisation for Economic Cooperation and Development
OLG  overlapping generations
OLS  ordinary least squares
PP  Phillips and Perron
PPP  purchasing power parity
PV  Perron and Vogelsang (1992a)
RET  Ricardian equivalence theorem
SBC  Schwarz Bayesian Criterion
SURE  seemingly unrelated regression
UK  United Kingdom
US  United States of America
VAR  vector autoregression
VECM vector error correction model
List of Notation

$A_t$  
central bank's assets, other than government debt

$\Delta$  
first-difference operator

$D(TB)_t$  
dummy variable equal to unity in the year after the break and zero otherwise

$D_d$  
seasonal dummies

$D_t$  
stock of interest-bearing real government debt

$d_t$  
stock of interest-bearing government debt, as ratio of GDP

$\delta_s$  
discount factor

$D_t^{CB}$  
government's debt in the central bank

$D_t^d$  
discounted value of real debt

$DU_t$  
dummy variable equal to unity in the years after the break and zero otherwise

$E_t$  
expectations operator, based on information available at time $t$

$GDP_t$  
gross domestic product, in nominal terms

$G_t$  
real primary government expenditures, including purchases of goods and services, and also transfers, but excluding interest expenditures

$g_t$  
primary government expenditures, as ratio of GDP

$i_t$  
nominal interest rate

$L$  
lag operator

$m_t$  
monetary base, as ratio of GDP

$M_t^N$  
nominal monetary base

$NS_t$  
net-of-interest government surplus, in real terms

$ns_t$  
net-of-interest government surplus, as ratio of GDP

$NS_t^d$  
discounted value of real net-of-interest surplus

$P_t$  
price level index

$\pi_t$  
inflation rate

$\pi_t^{CPI}$  
inflation rate, measured as the growth rate of the consumer price index

$\pi_t^e$  
extpected inflation rate

$r_t$  
real interest rate on government bonds

$\rho_t$  
real interest rate net of output growth

$r_t^{CPI}$  
ex-post real interest rate, deflated with the CPI

$S_t$  
real primary govern. surplus, excluding interest expenditures and seigniorage revenues

$s_t$  
primary government surplus, as ratio of GDP

$\sigma_t^{CF}$  
seigniorage revenues, cash-flow definition

$\sigma_t^{OC}$  
seigniorage revenues, opportunity cost definition

$T_B$  
date of the break

$TDEF_t$  
total gov. deficit, including interest expenditures and seigniorage revenues, in real terms

$tdef_t$  
total government deficit, as ratio of GDP

$TG_t$  
real total government expenditures, including interest expenditures

$tg_t$  
total government expenditures, as ratio of GDP

$tg^A_t$  
aggregate total expenditures, as ratio of GDP
tsₜ, aggregate total surplus, as ratio of GDP
Tₜ, real primary government revenues, excluding seigniorage
ᵣₜ, primary government revenues, as ratio of GDP
TTₜ, real total government revenues, including seigniorage revenues
ttₜ, total government revenues, as ratio of GDP
ttₜ, aggregate total revenues, as ratio of GDP
uᵣₜ,ₙ marginal rate of intertemporal substitution of consumption between periods t and t+n
ψₜ, real growth rate of output
Yₜ, gross domestic product, in real terms
1. INTRODUCTION

"Almost in the beginning was curiosity"
Isaac Asimov

The persistently large government deficits, and the resulting accumulation of debt in most developed countries since the mid seventies, raised significant concerns over the existence of long-run constraints on public borrowing, and the economic consequences of fiscal indiscipline. This has been one of the most controversial and discussed economic issues among academics and policymakers during the last two decades of the twentieth century.

The general view, namely at the political level, agrees that this tendency should be rapidly reversed. Accordingly, several countries have introduced in the last few years, or are planning to do so, institutional restrictions on fiscal policy, particularly on the size and financing of public deficits and debt. The European Union's fiscal convergence criteria are an adequate example, but similar measures have been adopted in the United States and Australia, among others.

Despite these recent efforts towards fiscal consolidation, the prevailing perspective is that the fiscal situation may deteriorate further in the future, given the expected demographic changes in most countries. An ageing population will impose increasing pressure on the already expensive welfare programs and especially on the unfunded social security systems, exerting a considerable strain on public finances over the next generations.

The main objective of this thesis is to investigate whether current fiscal policies are sustainable, that is, able to guarantee the government's solvency, and what are the consequences of unsustainability on monetization, inflation and interest rates. The empirical analysis is focused on a comparative study of six core members of the European Union (Belgium, France, Germany, Italy, the Netherlands and the United Kingdom), during the post-war period.

Although the most recent advances in European integration have implied fundamentally major changes in monetary policies, much of the academic and political debate has been
centred on fiscal policies. Sustainability and the fiscal effects on inflation and interest rates, have been a significant concern of the authorities in the process leading to the European Monetary Union:

"Unsustainable budgetary positions in a Member State, ultimately leading to either default or debt monetization, would be a major threat to the overall monetary stability. [...] In the medium-term, surveillance will have to correct possible tendencies for budget deficits to become too large as their interest rate cost is spread throughout the Union."

Commission of the European Communities (1990: p. 100)

The discussion over these questions is particularly important in a monetary union for several reasons. On the one hand, a deeper degree of economic and financial integration increases the probability that the effects of unsustainable fiscal policies in one country may spill-over to other member states, eventually threatening the stability of the whole union. On the other hand, the complete liberalisation of the financial markets, and the elimination of exchange rate risk, increases the internal mobility of goods, services and production factors, raising spending and tax competition and hence restricting national fiscal flexibility. This may be particularly problematic for a highly indebted country, where a significant fraction of public revenues is permanently reserved to debt service, reducing considerably its capacity to implement stabilisation policies and provide sufficient public goods. This could further deteriorate the fiscal situation, by jeopardising growth prospects and diverting the tax base.

However, the debate is not confined to the particular case of a monetary union. It is also a central concern at a more general level:

"This will permit us to think systematically about three sets of issues central to the [International Monetary] Fund's concerns. First, the issue of government solvency or sovereign default. Second, the issue of financial crowding-out of private saving and investment by government borrowing and, third, the monetary (and hence the inflation) implications of alternative fiscal-financial-monetary programs."

Willem Buiter (1995b: p. 2)

These three sets of issues have not yet been jointly addressed empirically in a unified framework. Although the empirical tests of sustainability of fiscal policies are motivated essentially by the potential economic consequences of unsustainability, highlighted in the public policy debate and supported in the conventional macroeconomic models, no
previous known study of sustainability has extended its scope to assess the consequences of different fiscal positions.

Sustainable fiscal policies are here defined as those which can continue unchanged into the future without violating the government's intertemporal budget constraint (IBC). This long-run constraint is basically an accounting identity requiring the outstanding stock of debt to be completely offset by the expected, in present value, sum of all future primary surpluses and money creation.

Until recently the budget constraint was assumed to hold permanently, as a restriction to the dynamic behaviour of fiscal policies, dispensing empirical testing. The IBC constitutes a central assumption in most macroeconomic models examining the effects of fiscal policies, in Barro's (1974) 'Ricardian equivalence theorem', Barro's (1979) 'tax smoothing', Sargent and Wallace's (1981) 'unpleasant monetarist arithmetic', and Lucas and Stokey's (1983) discussion of optimal taxation, among others.

Since the mid eighties, several approaches have been developed to support the theoretical debate with an empirical evaluation of the sustainability condition. Overall, the results show that in some situations the constraint cannot be assumed to hold as an uncontested equality. Two major strands have developed in parallel in the empirical literature. One was introduced by the seminal work of Hamilton and Flavin (1986), and it is focused on the restrictions imposed by the IBC on the long-run time series characteristics of the debt, deficit, and the various components of the government revenues and expenditures. Assuming that the statistical properties of the series remain unchanged, i.e., a stable data generating process, the (un)sustainability of the fiscal policy is assessed by examining the long-run univariate and multivariate stochastic properties of the variables.

The alternative approach to test sustainability requires, in addition to the IBC, a bounded debt-GDP ratio, and takes into account prospective developments of the relevant variables. The rationale for this 'stronger' approach, attributed to Blanchard et al. (1990), is the existence of limits to the government's capacity to raise revenues, reduce expenditures or place debt in the market. The proposed sustainability indicators measure the difference, for example, between the current tax rate and the 'sustainable tax rate' which guarantees a stable debt ratio, based on predictions over the evolution of the variables concerned. Different time horizons can be considered, ranging from the very
short-term to an infinite horizon. The longer the horizon, the more uncertain the availability and credibility of the projections. Ultimately, these indicators are as reliable as the forecasts upon which they are constructed.

The testing methodology adopted in this thesis favours the first approach, essentially to avoid the implicit normative criterion and the discretionary element involved in using projections of the variables. It also avoids the difficulty of evaluating feasible changes in policy. Unsustainability does not imply collapse but will necessitate a change of policy. However, the problems associated with a potentially ever-growing debt-GDP ratio should not be disregarded. If the investors react to an increasing ratio of debt by demanding higher interest rates or even by denying further lending, they can rapidly turn a sustainable into an unsustainable fiscal position. Therefore, the sequential testing strategy followed in the empirical application reconciles the first approach with the infinite horizon measure proposed by the second approach. The particular econometric methodology revolves around unit root and cointegration techniques, as suggested by the long-run equilibrium characteristics of the sustainability condition.

As emphasised above, the sustainability analysis will be complemented with the study of the fiscal impact on monetization, inflation and long-term interest rates. All these relationships between the variables can be derived and examined within a common theoretical construction, the IBC. The econometric methodology adopted to analyse these effects is centred on a vector autoregressive framework, adjusted according to the univariate statistical properties of the various time series being investigated. The major advantage of this multivariate approach is allowing the existence of feedback mechanisms between the variables, often neglected in the literature. The standard macroeconomic models of fiscal policy usually emphasise the effects from deficits and debt to the other variables and, therefore, most empirical studies employ a reduced-form equation derived from a complete model of the economy, where the fiscal variables are considered to be exogenous. However, inflation and interest rates are crucial variables in the budget dynamics, namely through their impact on interest costs, and hence on the deficit and debt accumulation processes.

The empirical evidence presented in this dissertation overwhelmingly suggests that current fiscal policies are unsustainable in most countries considered. Only Germany and
the Netherlands present sustainable policies in the long-run. However, it was not possible to reject in the latter the hypothesis of an unbounded growth of the debt-GDP ratio, indicating that the Dutch government may experience future difficulties marketing its debt. For all other countries, the current course of fiscal policies needs to be changed inevitably, either through a structural change in the budget, or a shock on the other macroeconomic variables such as growth, inflation or interest rates.

It was also found that unsustainable fiscal policies do not necessarily impose an excessive burden on monetary policy. High deficits do not imply increasing inflation in any country considered and, apart from Italy, do not generally cause higher interest rates. On the contrary, the tests suggest that high inflation and interest rates generally contribute to increase the deficit, mainly through their effects on the interest costs. There is however some evidence that, within liberalised financial markets, the real interest rate reacts to the level of the central government’s debt.

In sum, this thesis intends to make the following main contributions to the literature:

- analyse a long, but comparable, time period in a group of EU members. Previous studies usually employ samples starting in the early seventies, which may not provide a sufficiently long time span, given the characteristics of the formulated hypothesis and of the econometric methodology. Besides, the US fiscal situation has been by large the most closely scrutinised. No previous known study in this research area has analysed a longer time period for a group of EU countries;

- perform a comparative analysis of central and general government data, with the aim of assessing the influence of the regional budgets and, especially, the social security funds, considered the major source of public liabilities in the future;

- adopt a flexible approach to test the hypothesis of sustainability of fiscal policies, reconciling within a unified framework the two major strands of the empirical literature testing long-run sustainability;

- complement the sustainability analysis with the study of the fiscal effects on monetization, inflation and interest rates, using the same data set.
- apply the recent developments in econometric theory, specifically the tests of structural breaks, seasonal unit roots with breaks, panel data unit root tests, bounds-ARDL procedure, and LA-VAR Granger-causality;

The remainder of the thesis is organised as follows. Chapter two presents a comprehensive critical survey of the literature on sustainability of fiscal policies, analysing the alternative testing procedures, and emphasising their relative contributions and weaknesses. The chapter begins with a theoretical and analytical introduction of the IBC, and derives the restrictions it imposes on fiscal policy. The following sections contrast the two major approaches to test sustainability, the first requiring the stabilisation of the debt-GDP ratio and taking into account expected developments of the time series concerned, the second investigating the past statistical behaviour of the series.

Chapter three introduces the methodology, and describes the data sets employed in the empirical applications. The first section presents and justifies the methodology adopted to test sustainability, and very briefly reviews the main econometric techniques applied. A particular emphasis is justified in the case of the more recent techniques, some of which have not previously been employed in this research area. The second section of the chapter describes the main characteristics of the data sets, introduces the variables, the data sources, and the necessary pre-adjustments performed. It also offers a preliminary overview of the evolution of the most relevant variables involved in the IBC, advancing some possible explanations for the main common features in the data.

Chapter four performs a systematic empirical investigation of the long-run sustainability hypothesis, along the sequential testing strategy adopted. The same methodology is applied to three different data sets, containing annual and quarterly central government data, and annual general government data. Particular precautions are taken with the possible influence of structural breaks in the series. In addition, this chapter also considers the issue of how fiscal imbalances have been closed in the past, by increasing revenues, reducing expenditures, or a combination of both, and the related question of the most adequate policy to move out of an unsustainable position, considering the historical joint behaviour of the fiscal variables.

The next two chapters consider the fiscal policy effects on other macroeconomic variables. Chapter five discusses the effects on monetization, or seigniorage, and
inflation, probably the major concern with undisciplined fiscal policies. After briefly reviewing the literature on the subject, two major alternative measures of seigniorage are discussed and computed. The chapter examines separately the fiscal effects on seigniorage and on inflation, given the variety of channels of transmission between the variables. Finally, it also presents and estimates a simple public finance model of revenue collection, the only attempt throughout the thesis to incorporate optimality issues into the empirical discussion.

Chapter six investigates the influence of the fiscal variables, namely deficit and debt, on the long-term interest rate. The first part of the chapter surveys the theoretical and empirical literature on this question, intentionally contrasting the more conventional theories of a positive effect on the interest rate, with the alternative opposing views, notably the Ricardian equivalence theorem. The chapter investigates this issue thoroughly, with the purpose of increasing the robustness of the conclusions, and explores different perspectives of the same basic question.

Finally, chapter seven summarises the most significant findings of the thesis, highlights the main contributions and weaknesses, and suggests potential future directions of research in this area.
2. THE IBC AND ALTERNATIVE SUSTAINABILITY INDICATORS

2.1. Introduction

The academic debate over the sustainability of a fiscal policy is closely connected with the government's intertemporal budget constraint (IBC). According to this constraint, the current stock of debt must be exactly matched by the sum, in present value, of all future net-of-interest surpluses. This requires the complementary restriction that the discounted value of the debt-GDP ratio must equal zero in the limit.

Several distinctive tests have been proposed to investigate this long-term notion of fiscal policy sustainability with a transversality condition. One major difference between the alternative approaches rests on the assumptions concerning the statistical properties of the interest rate, a pivotal variable in the dynamics of the IBC. Section 2.4 reviews the different procedures thoroughly, explores their main comparative advantages and gaps, and presents the results of some empirical tests based on these distinct approaches.

For some authors, however, this notion of sustainability is not sufficiently 'strong' to be reassuring. The condition requiring the discounted value of the debt ratio to approximate zero only in the limit, in the indefinite future, allows the actual value of this ratio to grow without bounds. They argue that this condition is not realistic, given for example the natural maximum limits on the capacity of the state to tax its residents.

In this sense, satisfying the IBC would be a necessary but not sufficient condition to guarantee the sustainability of fiscal policies. An additional restriction should be imposed, usually referred to as the 'collateral constraint'. It requires the stabilisation of the debt-GDP ratio in a more or less distant, but finite, horizon. In its most basic form, this condition tests whether the current primary surpluses are sufficient ('sustainable') to ensure the immediate stabilisation of the current level of public debt.
Less restrictively, these 'stronger' tests may also be used to assess the capacity of convergence to a certain level of the debt ratio, within a certain number of periods, stabilising at that 'optimal' level afterwards. This is basically the idea behind the Maastricht fiscal convergence criteria, for example. In reality, these criteria introduced a new concept of "sustainability of the government financial position" (Art. 109J-1.), based on maximum arbitrary ceilings for the fiscal variables. Von Hagen (1998) labelled this new concept as MTD, 'maximum tolerable debt'.

A more elaborated version of these stronger sustainability tests, requiring the stabilisation of the debt ratio, was proposed by Blanchard (1990). It has the main advantage of including in the analysis projections of future values of the relevant variables. It is also constructed in a way that allows the quantification of the precise fiscal adjustment necessary to achieve sustainability. The major problem with this forward-looking test is its high dependence on the quality of these projections.

This chapter is structured as follows. Section 2.2 introduces the notion of intertemporal budget constraint, and its main implications for the different empirical tests of sustainability. Section 2.3 examines the tests of the stronger measure of sustainability, with a collateral constraint imposing a limit on the debt-GDP ratio. In section 2.4 this collateral constraint is abandoned, and instead a less strict condition is imposed. The testing approaches surveyed in this section are based on the analysis of the long-term stochastic properties of the variables involved in the IBC. Finally, section 2.5 recapitulates the main ideas of the chapter, and summarises some of the empirical results reported in the literature, with a special emphasis on those concerning the European Union.

2.2. The Intertemporal Budget Constraint

The literature on the sustainability of fiscal policies unavoidably adopts the IBC as a starting point of the analysis. This constraint relates the joint and interdependent evolution through time of the government's debt and deficit, two central indicators of a country's fiscal policy. The formalisation of the government's intertemporal constraint is
analytically derived from a forward solution of the one-period budget identity, which can be expressed, with all the terms measured in real values, as

\[ D_t = (1 + r_t)D_{t-1} + G_t - T_t - \frac{\Delta M_t^N}{P_t}, \]

where \( D_t \) is the stock of interest bearing real public debt outstanding at the end of the period, and \( r_t \) is the ex post real interest rate. For purposes of simplicity, this debt can be assumed to take the form of one-period single coupon bonds, issued at par. Otherwise, the interest rate must be interpreted as an average rate of return on the global stock of debt, including the coupon payments and the change in the market value of the asset. The variables \( G_t \) and \( T_t \) represent, respectively, real primary government expenditures, i.e., excluding interest payments on the public debt, and real primary government revenues, excluding seigniorage. Both variables are net of transfers. The last term designates the real revenues from money creation, where \( M_t^N \) stands for the monetary base in nominal terms, deflated by the price level \( P_t \). All stocks are measured at the end of the period and \( \Delta \) is the usual first-difference operator.

Defining \( S_t \) as the real primary balance \((T_t - G_t)\), i.e., a surplus if positive and a deficit if negative, and rearranging, yields the cash-flow equation

\[ \Delta D_t = r_t D_{t-1} - S_t - \frac{\Delta M_t^N}{P_t}, \]

Solving (2.2) for \( n \) periods forward by recursive substitution, and letting \( E_t \) denote the expectations' operator based on the information available at time \( t \), produces the government's present value, or intertemporal, budget identity

---

1 This is a very concise expression of the public accounts. Some authors (Buiter, 1983, and Corsetti, 1991, for example) use an expanded version of this identity by disaggregating some of its variables (e.g., total government spending decomposed into consumption, replacement investment and net investment; bonds divided by maturity and country of placement; taxes separated from the revenues accruing from selling of public assets, income from public sector capital and natural resource property rights, etc.). However, the data required for an empirical evaluation of these more elaborated equations is not available for the whole time period and for all countries, demanding restrictive assumptions, simplifications and generalisations (see, for example, Blejer and Cheasty, 1991).

2 It was assumed that \( P_t \), usually given by the GDP deflator, is the appropriate deflator for all variables in equation (2.1). Otherwise, the equality symbol in this equation would not stand.

3 Solving (2.2) backwards would depict the present value of public debt as the result of the history of all budget deficits incurred previously, an accounting identity not useful for testing purposes. Therefore, all
The gross compound interest rates, or discount factors, are given by

\[(2.4) \prod_{i=1}^{n} \left( \frac{1}{1+r_{it+j}} \right) = \delta_{t,n}, \]

henceforth represented by \( \delta_{t,n} \) to simplify the notation. For practical reasons, it is sometimes assumed in the empirical literature that the interest rate is constant, in which case the discount factor reduces to \( \delta_{t,n}=(1+r)^{-n} \).

The present value budget identity should also hold for an infinite horizon, yielding

\[(2.5) \quad D_i = \lim_{n \to \infty} E_i \delta_{t,n} D_{t+n} + \sum_{j=1}^{n} E_i \delta_{t,j} \left( S_{t+j} + \frac{\Delta M_{t+j}^N}{P_{t+j}} \right). \]

To turn the accounting identity (2.5) into a solvency constraint, it must be assumed that the first term on the right-hand side equals zero, ruling out everlasting 'Ponzi finance schemes' - the government cannot forever pay the interest on its outstanding debt simply by continuously rolling it over totally with additional borrowing. It is assumed that lenders do not tolerate this situation indefinitely, because they could increase their consumption by not buying bonds, and therefore

\[(2.6) \quad \lim_{n \to \infty} E_i \delta_{t,n} D_{t+n} = 0. \]
This terminal, or 'transversality condition' ensures that the discounted value of the total government debt approaches zero in the long run. The transversality condition does not imply that the debt has to go to zero, nor even towards a finite value. It simply states that it must grow at a rate strictly less than the rate of interest - which will happen as long as the government does not present primary deficits (McCallum, 1984) or these are completely paid for by base money creation (Mitchell, 1988).

If the limit in (2.6) is positive, the current fiscal policy is unsustainable and must be changed in order to satisfy the IBC. If the limit reveals a negative sign it indicates a situation of supersolvency, where the government, in the limit, is a net creditor, not a very plausible situation (Corsetti, 1991). A negative expected discounted debt would mean that other economic agents were running Ponzi games against the government, implying the maintenance of higher taxes, or lower public spending, than necessary to service the debt.

If (2.6) holds exactly, equation (2.5) becomes the government’s present value borrowing constraint, or intertemporal budget constraint (IBC),

\[
(2.7) \quad D_t = \sum_{j=1}^{\infty} E_t \delta_{t,j} \left( S_{t+j} + \frac{\Delta M_{t+j}}{P_t} \right).
\]

The present value of the government’s debt must equal the sum of the discounted values of all expected primary surpluses and base money creation. Equation (2.7) discloses the intra- and intertemporal trade-off between two different sources of government revenues, taxes and money creation. It exposes, for example, the Sargent and Wallace’s (1981) well-known proposition that sooner or later permanent primary deficits will necessarily be monetized. On the other hand, ignoring the possibility of monetary revenues, it also depicts the inevitable choice between debt and tax financing of the government’s stream of expenditures. The effects of these choices on the economy will be examined in chapters 5 and 6.

This equation can be alternatively interpreted as representing the evaluation of the equilibrium price of the government bonds by the discounted stream of its future net

---

8 The idea of a terminal condition may seem contradictory with the conventional view that governments are infinitely lived. However, this is not relevant in practice since, considering a positive discount rate, the
revenues. The equivalence between both interpretations of (2.7) is plausible since the par
and the market value of the bonds are the same at the date of emission. As will be noted
in section 2.4, the long-run tests of sustainability were inspired by this analogy with the
asset pricing theory.

For reasons of international and intertemporal comparisons, it is sometimes convenient to
define these fiscal variables as ratios to GDP, with the real interest rate net of output
growth $\psi$ and symbolised by $p$. Representing with lower case characters all the other
corresponding variables that have been divided by real GDP, equations (2.1) and (2.2)
become

\begin{equation}
(2.8) \quad d_t = \frac{(1+r_g)}{(1+\psi_t)}d_{t-1} + g_t - t_t - \Delta m_t - m_t \left[1 - \frac{1}{(1+\pi_t)(1+\psi_t)}\right], \text{ and}
\end{equation}

\begin{equation}
(2.9) \quad \Delta d_t = \rho_t d_{t-1} - s_t - \Delta m_t - m_t \left[1 - \frac{1}{(1+\pi_t)(1+\psi_t)}\right],
\end{equation}

where $\rho_t = (r_g, \psi_t)/(1+\psi_t)$. The path of debt is a negative function of the inflation and
growth rates, either through a devaluation of the accumulated amount of debt, or through
a devaluation of the monetary base.

This presentation in ratios is consistent with the usual indicators employed to evaluate the
performance of fiscal policies (in the Maastricht Treaty, for example), and it will be
useful later to distinguish between sustainability in the broad sense and stability of the
debt-GDP ratio. The other important advantage is that it removes from the tests potential
effects of nonstationarity in inflation and real GDP growth (Makrydakis et al., 1996: p.
6). However, the argument that tests based on (2.8) rather than on (2.1) are more correct
on a growing economy (as maintained, for example, in Hakkio and Rush, 1991a) is not
accurate, since the IBC derived from any of these equations is formally equivalent and
ultimately it is not dependent on the growth rate of the economy (a formal proof can be
found, for example, in Cuddington, 1997: p. 12).

Equation (2.9) is a difference equation with its dynamics exclusively dependent on the
values of the growth-corrected real interest rate (Table 2.1).

weight of the fiscal variables in the distant future tends to zero.
Table 2.1: The dynamics of the debt-GDP ratio

\[ p < 0 \Rightarrow \text{if the net-of-interest surplus is a stationary process with zero mean, then the debt-GDP ratio converges to zero. This is a sufficient, albeit not necessary, condition to obey the IBC. In fact, the government is not subject to a real constraint, since it can issue new debt indefinitely without needing to repay it. When the growth rate of the economy is sufficiently high, the stability of the debt ratio is compatible even with continuous primary deficits.} \]

Equation (2.9) shows that the growth rate of the economy can finance the whole interest service plus a part of the net-of-interest deficit;

\[ p = 0 \Rightarrow \text{if the net-of-interest surplus is stationary with mean zero, then the debt-GDP ratio is expected to remain stable in the long-run. Stability of the debt ratio could be achieved with a permanent net-of-interest budget balance, although this is not necessary since it just has to be balanced on average;} \]

\[ p > 0 \Rightarrow \text{equation (2.9) becomes an unstable deterministic difference equation. If the government continuously runs primary budget deficits, the debt to output ratio becomes explosive. In this case a solvency, transversality condition is needed in order to limit the growth of the public debt. The stability of the debt ratio requires necessarily a primary budget surplus.} \]

Assuming, for simplicity, a constant net-of-interest deficit and discount rate, the solution of the difference equation (2.9) indicates that the ratio of debt to GDP converges to a value given by the ratio of the net-of-interest deficit and the growth-corrected real interest rate. In this situation, Buiter (1985: fn. 10) claims that the governments are playing 'honest' Ponzi games. Strictly speaking, there is no need for primary surpluses, as long as condition (2.7) can be met exclusively with seigniorage revenues. If these are sufficiently large, continuous primary deficits are still feasible (see, for example, Mitchell, 1988).

In the empirical literature, it is usually assumed that condition \( p > 0 \) is expected to hold. In fact, although during the 1970's most countries have maintained a growth rate of the economy above the real interest rate on public debt, it is difficult to suppose, in the long run, a stable economic growth without a sufficiently high capital remuneration (McCallum, 1984). Barro (1985) dismisses the hypothesis \( p < 0 \) because it is inconsistent with maximising behaviour by the economic agents, since it signals an unexploited profit opportunity. Furthermore, if the interest rate is not independent of the fiscal stance, it can be argued that even if in a certain moment in time the growth rate of the economy exceeds the real interest rate, the pressure of growing amounts of debt may eventually reverse the situation (as will be shown in chapter 6). Therefore, even if in some periods

Assuming, for simplicity, a constant net-of-interest deficit and discount rate, the solution of the difference equation (2.9) indicates that the ratio of debt to GDP converges to a value given by the ratio of the net-of-interest deficit and the growth-corrected real interest rate. In this situation, Buiter (1985: fn. 10) claims that the governments are playing 'honest' Ponzi games. Strictly speaking, there is no need for primary surpluses, as long as condition (2.7) can be met exclusively with seigniorage revenues. If these are sufficiently large, continuous primary deficits are still feasible (see, for example, Mitchell, 1988).

In the empirical literature, it is usually assumed that condition \( p > 0 \) is expected to hold. In fact, although during the 1970's most countries have maintained a growth rate of the economy above the real interest rate on public debt, it is difficult to suppose, in the long run, a stable economic growth without a sufficiently high capital remuneration (McCallum, 1984). Barro (1985) dismisses the hypothesis \( p < 0 \) because it is inconsistent with maximising behaviour by the economic agents, since it signals an unexploited profit opportunity. Furthermore, if the interest rate is not independent of the fiscal stance, it can be argued that even if in a certain moment in time the growth rate of the economy exceeds the real interest rate, the pressure of growing amounts of debt may eventually reverse the situation (as will be shown in chapter 6). Therefore, even if in some periods

Mainly because the decline in investment due to rising input prices, especially energy, and high inflation rates caused negative ex post real interest rates (Wilcox, 1983).
\( \rho < 0 \), it is not theoretically conceivable that this situation may last indefinitely. Consequently, it seems reasonable to consider the 1970s as a singular period and that the appropriate expectation is for \( \rho > 0 \) (Horstmann and Schneider, 1994: p. 360).

In addition, the hypothesis of a positive interest/growth rate differential is usually considered necessary for the notion of 'dynamic efficiency', in a nonstochastic economic model of overlapping generations like the one of Diamond (1965).\(^{10}\) An economy where \( \rho < 0 \) is dynamically inefficient, since it is possible to have a 'Pareto improvement' for the different generations. In a situation of dynamic inefficiency, where the economy has too much capital and a decumulation would improve welfare, the government can play a Ponzi game for as long as the lenders are willing to hold the outstanding stock of debt.

Conversely, an economy with under-accumulation of capital, corresponding to a marginal product of capital higher than the growth rate of output, is dynamically efficient. It is not possible to achieve the optimum (or 'golden rule of accumulation', a concept due to Phelps, 1961, where \( r = \psi \) and therefore \( \rho \) is null) without reducing the utility of at least one generation. In terms of equation (2.9), if a constant interest and growth rates are assumed, the condition \( \rho > 0 \) must always be satisfied. However, if those rates are allowed to vary, the condition becomes that \[ \sum_{j=1}^{n} \prod_{i=1}^{t} \left( \frac{1 + \psi_{j,i}}{1 + r_{i,j}} \right) \] must converge to a finite limit.

In fact, Abel et al. (1989) found evidence that all the seven economies they tested are dynamically efficient. However, they claim that 'dynamic efficiency' cannot be assessed simply by comparing the interest rate with the growth rate of the economy, once one allows for risk averse investors and therefore a stochastic economy. The main theoretical finding of their paper is that the economy will be dynamically efficient only if the cashflow generated by the stock of capital exceeds the value of the new investment. In a deterministic economy, with an equilibrium growth trajectory (characterised, among other aspects, by a constant capital-output ratio), the previous reasoning is correct. The income from capital equals the product of the interest rate times the capital stock, while the investment equals the product of the real growth rate of the economy times the capital

---

\(^{10}\) See, for example, Blanchard and Fisher (1989).
stock. Therefore, the capital income exceeds investment only if the interest rate is higher than the growth rate of the economy.

However, in the presence of uncertainty these conditions are not equivalent, and dynamic efficiency is compatible with a negative, growth corrected, real interest rate. This finding induced Bohn (1995) to criticise and reject all sustainability tests grounded on a deterministic economy, as will be examined later in this chapter.

With the redefined fiscal variables as ratios to GDP, the transversality condition and the IBC, equations (2.6) and (2.7), are now expressed as

\[ \lim_{n \to \infty} E_t \delta_{t,n} d_{t+n} = 0, \]  

and

\[ d_t = \sum_{j=1}^{T} E_t \delta_{t,j} \left[ s_{t,j} + \Delta m_{t,j} + m_{t,j-1} \left[ 1 - \frac{1}{(1 + \pi_{t,j})(1 + \psi_{t,j})} \right] \right], \]

with the discount factor of the model expressed as ratios to GDP being

\[ \delta_{t,j} = \prod_{k=1}^{j} (1 + \rho_{t,k})^{-1}. \]

These two equations suggest a simple theoretical definition for the concept of solvency: a government is considered solvent if its fiscal policy generates a sequence of debt and deficits in such a way that the present-value borrowing constraint holds. Conversely, with an insolvent government, the evolution of the fiscal policy does not cause the expectation of the discounted value of the debt to go to zero in the limit.

The concept of solvency is basically similar to the one faced by any individual or a firm. However, a considerable difference between the public and the private sector is that the former usually possesses a considerable control over its future revenues, either by enforcing a rise in taxes, for example, or by increasing the seigniorage revenues. Another important difference is that individuals have finite lifetimes whereas governments as a whole do not. In reality, "only if there exists no economically and politically feasible set of tax, spending and seigniorage plans that permit the existing stock of debt to be

\[ 11 \text{ Assuming, for ease of exposition, constant real interest and growth rates, it is simple to demonstrate the equivalence between equations (2.6) and (2.10). Knowing that } Y_{t+n} = Y_t (1 + \psi)^n, \text{ where } Y_t \text{ stands for real GDP, and } d_t = D_t Y_t^\psi, \text{ it follows that } d_{t+n}(1+\rho)^n = d_{t+n}(1+\psi)(1+r)^n = D_{t+n} Y_t^\psi (1+\psi)(1+r)^n, \text{ from where it is shown that } \lim_{n \to \infty} \delta_{t,n} D_{t+n} = 0. \]
serviced, can one truly speak of insolvency" (Buiter, 1990: p. 147). The transversality condition (2.10) only restricts the evolution of the discounted stock of debt asymptotically, while equation (2.11) does not restrict the government fiscal policy over any finite time interval. "The intertemporal budget constraint imposes restrictions only on the long-run relationship between expenditures and revenues, so that almost any short-run deficit path is consistent with a budget balanced in present value terms" (Trehan and Walsh, 1988: p. 425).

This makes it difficult to assess in practice the hypothesis that the government intertemporal budget is balanced, mainly because it depends on unobservable elements such as the present value of all net-of-interest surpluses and the interest and growth rates. It is necessary to impose some auxiliary assumptions in order to allow the testing of sustainability as a restriction to fiscal policy.

Two types of additional assumptions may be imposed, allowing two markedly different approaches to the testing of sustainability. The first approach is to consider the existence of a 'collateral constraint', requiring the stabilisation of the debt-GDP ratio. It implicitly or explicitly acknowledges the existence of certain limits to the medium-term evolution of some particular fiscal variables: social and economic limits to taxation and public spending or political self-imposed limits on debt and/or deficits, for example. In these circumstances the requirements for a sustainable fiscal policy will be even stricter, and the average growth rate of real debt (debt ratio) must be bounded by a number smaller than \( r (\rho) \). The next section will analyse the implications of adopting this first assumption.

---

12 The expressions 'fiscal policy sustainability' and 'government solvency' are often used indistinguishably in the literature. However, there are subtle theoretical differences between the two concepts, although both relate to the IBC. This thesis does not intend to test whether a government is solvent, it will investigate whether the current fiscal policy is sustainable. Theoretically, any government is solvent since it has a direct control over its own revenues. It is sufficient to assume that, somewhere in the future, it will provide the necessary primary surpluses to compensate for the accumulated deficits. The concept of sustainability is more restrictive, since it is assessed assuming no change in the relevant macroeconomic variables.

13 If the debt grows at a rate lower than the interest rate, but higher than the growth rate of GDP \((\psi < \Delta D/D, \lambda < \rho)\), the debt-GDP ratio will grow without bound.

14 There is no terminology generally accepted in the literature to distinguish these two approaches. Here they will be identified as 'sustainability with a collateral constraint' and 'sustainability with a transversality condition'.

15 "Intuitively, the concept of sustainability seems to be associated with both time-invariance and feasibility aspects" (Bohn, 1991b: p. 586).
The second major conceptual approach to test sustainability is to consider the long-term viability of the current fiscal policy and macroeconomic environment. Assuming that the present statistical properties of the relevant time series continue unchanged into the future, will the IBC and transversality conditions be satisfied, or is there a need for a fiscal policy change? This question will be examined in section 2.4.

In sum, both approaches are based on the IBC. The fundamental difference between them is the way they deal with the term \( \delta_t d_{t+n} \) in the intertemporal budget identity. While the first approach assumes either that \( d_{t+n} = d_t \) (stabilisation of the debt ratio at its current value), or that \( d_{t+n} = d^* \) (stabilisation at a certain 'optimum' level), the second type of tests assumes that this term approaches zero as \( n \) increases (\( \lim_{n \to \infty} E_t \delta_t d_{t+n} = 0 \)). The first approach sets a limit on the level of debt relative to GDP while the second sets a limit on its growth rate.

Another important difference is that while one is based on current and future expected values of the relevant variables, the other looks exclusively at current and past observations. The main problem of the first approach is the potential fallibility of the projections. The second, on the other hand, is vulnerable to future alterations in the structure of the time series.

2.3. Sustainability with a collateral constraint

The transversality condition (2.6) allows the debt-GDP ratio to grow without bound, which will happen as long as the growth rate of debt remains below the interest rate, but above the growth rate of the economy. This boundless growth of the debt ratio is considered by some authors as an unrealistic hypothesis. Blanchard (1984) and Spaventa (1987) argue that there is a maximum tax burden and a minimum level of public expenditures socially accepted. Considering specifically the redistributive effects, Keynes claims in 1923 that "the active and working elements in no community, ancient or modern, will consent to hand over to the rentier of bond-holding class more than a certain proportion of the fruits of their work" (Keynes, 1971: p. 54). Buiter and Patel (1992: p. 186) argue that only finite debt-GDP ratios are feasible, given for example the existence
of "deadweight losses or collection costs". On the other hand, Kritchell et al. (1996) point to imperfect capital markets, and Sargent and Wallace (1981) to limits on the private demand for bonds. As the debt ratio grows continuously, the government will be faced with an increasing difficulty to place its debt on the market, since the investors at some point will start considering the probability of default. An increasing risk premium is demanded on the interest rate of government bonds and, eventually, further borrowing may be denied.

To avoid all these theoretical problems an additional constraint is required, sometimes referred to as the 'collateral constraint', ruling out such possibility. This collateral constraint may demand a change in the fiscal policy even if the long-term sustainability conditions implied by equations (2.10) and (2.11) have been satisfied. This is why Buiter and Patel (1992) regard this stabilisation of the debt-GDP ratio as a stricter sustainability hypothesis, by comparison with the 'weak' hypothesis imposed solely by those equations.

### 2.3.1. Stabilisation of the current debt-GDP ratio

The stabilisation of the debt-GDP ratio at its current level implies that the left-hand side of equation (2.9) equals zero, yielding

\[
\rho_t d_t = s_t + \Delta m_t + m_{t-1} \left[ 1 - \frac{1}{(1 + \pi_t)(1 + \psi_t)} \right].
\]

This solution of the difference equation (2.9) can equally be derived from the IBC, equation (2.11), considering constant, steady state values for the debt and deficit exclusive of interest payments.

Condition (2.12) requires the government to run net-of-interest surpluses to compensate for debt servicing costs. The test of sustainability consists merely in multiplying the current debt ratio by the growth-corrected real interest rate, obtaining the net-of-interest surplus necessary to stabilise \(d\) at its current level. Comparing this value with the actual net-of-interest surplus, it is possible to assess the sustainability of the current fiscal policy. This 'permanent primary gap' was first proposed in Buiter (1983 and 1985).

Using this simple criterion of sustainability for a large sample of developed countries, an empirical test in OECD (1990) concludes that fiscal policies appear to be widely
sustainable, with the only exceptions of Greece, Italy and the Netherlands. Blanchard et al. (1990) add to this group Norway and Spain. Heinemann (1993), on the contrary, claims that with the exception of Belgium, Ireland and Luxembourg, all other nine EC members at the time follow unsustainable policies. Kremers (1989) finds evidence that the US fiscal policy stabilised the debt-GDP ratio until 1981 but not afterwards.

The rationale for considering the current debt ratio as a sustainable value is the assumption that if no default has occurred, this value cannot be considered to be beyond the country's ability to pay. However, instead of aiming at the immediate stabilisation of the debt ratio at precisely its present level, the objective of the authorities may be to stabilise it at a certain optimal level in a pre-determined future date. The Commission of the EC (1990), in the report 'One market, one money' claims that debt-GDP ratios above 100% should be stabilised. Many countries are now facing voluntary medium term upper limits on government's debt and deficits. In the United States and Australia, for example, there have been proposals to balance the budget by the beginning of the next century. In the European Union, the Maastricht's convergence criteria impose maximum limits on the deficit and debt, as a necessary condition to participate in the final stage of monetary union.\(^16\)

The analysis of sustainability with a collateral constraint is therefore particularly relevant when there is a medium term target to achieve, in terms of a particular level of debt or deficit. The objective is to evaluate whether the current economic policy will allow that target to be achieved in time or not. For example, suppose that the aim is to converge to a level of debt-GDP ratio \(d^*\) within \(n\) periods, and stabilise it at this optimal level afterwards. From equations (2.10) and (2.11), for a finite horizon, and defining \(d_{t+n}^*\) as the desired target-value for the debt-GDP ratio in period \(t+n\), the intertemporal budget identity becomes

\[
d_i - E_i \delta_{i,n} d_{i+n}^* = E_i \sum_{j=1}^{\infty} \delta_{i,j} \left[ s_{i,j} + \Delta m_{i,j} + m_{i,j-1} \left( 1 - \frac{1}{(1 + \pi_{i,j} \gamma (1 + \psi_{i,j}))} \right) \right].
\]

This equation shows that the difference between the actual debt ratio and its expected discounted value \(n\) periods ahead, must equal the sum of the discounted primary

\(^{16}\) The next section examines the particular case of the Maastricht's convergence criteria.
surpluses and base money creation until period \( t+n \). The total value of these net-of-interest surpluses will depend on the amplitude of the required debt reduction, the time interval allowed, and also the value of the discount factor, more specifically the real interest and growth rates.

Assuming a constant growth and real interest rates in equation (2.13), it is possible to compute the constant net-of-interest surplus necessary during all the next \( n \) periods in order to achieve the desired objective for the debt ratio

\[
(2.14) \quad [d_t - (1 + \rho)^{-n} d_{t+n}^*] \left[ \frac{\rho(1 + \rho)^{n-1}}{(1 + \rho)^n - 1} \right] = s_t + \Delta m_t + m_{t-1} \left[ 1 - \frac{1}{(1 + \pi_t)(1 + \psi_t)} \right].
\]

After period \( t+n \), stabilisation of the stock of debt at the level \( d^* \) requires a constant flow of net-of-interest surpluses given now by equation (2.12).

2.3.2. Sustainability in the Maastricht Treaty and the Stability Pact

In 1991, the members of the European Community signed the Treaty of the European Union, commonly referred to as the 'Maastricht Treaty', where some rules were established as eliminatory conditions for the countries wishing to participate in the final stage of the European Economic and Monetary Union. Among others, two convergence criteria were then agreed in terms of the behaviour of fiscal policies: a maximum limit of 60% of GDP for the stock of public debt, and a ceiling of 3% for the total deficit relative to GDP.

Since then, several studies have tried to depict the theoretical justifications for such precise and apparently arbitrary numbers. This section presents some of these possible justifications, and considers how these rules can be examined in the light of the sustainability indicators analysed in the previous section.17

---

17 Woodford (1998) designates this 60% ceiling as the 'transversality condition'.

21
The limit of 60% for the ratio debt-GDP corresponds, approximately, to the average value of this ratio for the whole European Union in the years preceding the Treaty. The main criticism to the choice of that (or any) particular number as the optimum for all member states, is that the average of 60% conceals important differences in national economic structures and initial conditions, demanding unequal convergence efforts.

The same simple justification cannot however be applied to the ratio deficit-GDP, whose average value for the European Union has been usually above 3%. Two possible theoretical explanations for the choice of this precise value can therefore be advanced. The first attempt to justify the choice of 3% as a universal value of convergence is based on the so called ‘golden rule of public finances’: balance the current budget and borrow only to finance capital expenditures. Theoretically, if the discounted value of the return flow from this capital is equal to the initial investment cost, this type of expenditure would not have any consequences in terms of an intertemporal analysis. Coincidentally, the capital expenditures in the EU averaged precisely 3.02% of GDP during 1974-89, and 3% in 1990 and 1991 (values from Buiter et al., 1993).

This explanation has been highly contested mainly because it does not take into account the effects of inflation on the depreciation of public debt. Suppose a steady state equilibrium in which the two fiscal criteria are exactly verified with the maximum values allowed, and public investment is 3% of GDP. If this hypothetical economy presents an inflation rate of 2%, the government’s real savings will be approximately 1.2% of GDP, corresponding to a decline in the real value of public debt. This amount would be even higher if it also included the revenues from the increase in real balances and from the reduction in the real value on nominal base money. In this situation, public investment will be financed by borrowing in only around 1.8% of GDP, showing that the ‘golden

---

18 The Maastricht European Council have reached agreement on the draft of the ‘Treaty on EU’ in December 1991. Data from the Commission services available in November 1991 indicates the following values for the EU’s average debt as a per cent of GDP for the years between 1985 and 1991 (forecast): 59.1, 60.0, 61.5, 61.0, 60.3, 61.8 (Commission of the European Communities, 1991).

19 Originally applied to the German Länder. The Maastricht treaty refers to this ‘golden rule’, although it does not mention it by this expression (art. 104.-C, no. 3).

20 Other criticisms include the practical problems of distinction between current and capital spending (Begg et al., 1991), the economic arguments in favour of borrowing to finance public consumption expenditures, and the fact that, in practice, it is not always possible to assume that the discounted flow of future receipts from these investments equals its initial cost (see, for example, Fisher and Easterly, 1990).
rule' could be satisfied even with higher reference values for the debt and deficit. The Maastricht fiscal convergence criteria seem to be even more severe than the golden rule.

The second possible justification for this particular deficit limit, is to consider the 3% as the value compatible with a long-run, steady-state equilibrium (Corsetti and Roubini, 1992), the guarantee of a stable debt-GDP ratio of 60%. This can be shown analytically starting from the stabilisation condition in equation (2.12), requiring that the primary surplus should equal, in every period, the debt ratio multiplied by the growth-adjusted real interest rate.

This equation must however be slightly modified since the Maastricht’s criteria were set in terms of the total, not the primary surplus. Since \((1+\rho) = (1+i)[(1+\pi)(1+\psi)]\), where \(i\) is the nominal interest rate, taking logarithms to this expression, it can be approximated by \([\rho = i - (\pi + \psi)]\). Substituting in (2.11) yields

\[
(\pi + \psi) d_t = \rho_t \left\{ s_t + \Delta n_t + m_{t-1} \left[ 1 - \frac{1}{(1+\pi)(1+\psi)} \right] \right\}.
\]

The right-hand side of (2.15) is the government’s total deficit as a percentage of GDP \((tdef)\), as defined in the Maastricht’s Treaty,

\[
tdef = (\pi + \psi) d.
\]

Considering a nominal growth rate of 5% (admitting approximately, an inflation rate of 2% and a real growth rate of 3%),\(^{21}\) equation (2.16) shows that the limit of 60% chosen for the ratio debt-GDP must be matched by a value of 3% for the ratio total deficit-GDP. One useful rule of thumb to guarantee sustainability in a steady-state equilibrium would then be

\[
TDEF / D \leq (\pi + \psi)
\]

One interesting conclusion derived from this expression is that both fiscal criteria can not be chosen independently, and consequently only one of the Maastricht’s fiscal criteria is necessary for sustainability purposes. Winckler et al. (1998: p. 273) go further, claiming

---

\(^{21}\) These are the values frequently used as projections in most official European Union’s studies. The report ‘One Market, One Money’ (Commission of the European Communities, 1990), for example, analyses the
that “the Maastricht criteria are neither necessary nor sufficient for the sustainability of a deficit-debt pair”.

The ‘Stability and Growth Pact’, signed in Amsterdam in 1997, intends to reinforce some of the fiscal restrictions presented in Maastricht. In short, the Stability Pact determines the need for balanced budgets over the economic cycle, allowing the automatic stabilisers to work during cyclical downturns, i.e., the 3% ceiling suggested in Maastricht should be seen only as an exception. The Pact also includes fines for those countries with deficits exceeding the 3% limit. As may be observed from the simple analytical demonstration above, the Stability Pact is even more rigorous than the Maastricht Treaty. By requiring an average balanced budget, it also suggests that the debt-GDP ratio should approach zero in the long-run.

2.3.3. Blanchard’s sustainability indicators

Blanchard (1990) argues that the above simple measure of sustainability, based on the stability of the debt-GDP ratio in the steady-state, has two important main limitations. In the first place, it assumes the continuation into the future of the present fiscal conditions, not allowing for predictable developments in fiscal policy or in the other relevant macroeconomic variables. A certain fiscal policy, seemingly unsustainable according to this measure, may in reality be sustainable because the necessary correction has already been planned for the future. Conversely, an apparent sustainable policy may become unsustainable if, for example, an expected increase in expenditures due to unfunded social security schemes is included in the calculations.

A second limitation of the above sustainability tests is that they do not provide any quantitative indication about the magnitude of the necessary policy change. Although able to send signals to the authorities, indicating if they should continue with the current fiscal stance, these tests may not be very decisive for operational policy purposes, in terms of the government’s decisions on fiscal policy.

To overcome these limitations, Blanchard (1990) suggests a ‘new set of fiscal indicators’ to evaluate the sustainability of fiscal policy. The first limitation is solved by considering question of the government’s solvency assuming these values, except for a real growth rate of 3.5% for the
current projections for the relevant variables, while the second is overcomed by estimating the precise value of the fiscal adjustment required to achieve sustainability.\textsuperscript{22}

In theory, the necessary adjustments can be achieved through a cut in public spending, an increase in taxes or resorting to monetary financing. In practice, the empirical studies usually employ the tax rate as the endogeneous variable, assuming that "taxes are more likely to be the factor which is adjusted" (Blanchard, 1990: p. 13).\textsuperscript{23}

In practice, the empirical studies usually employ the tax rate as the endogenous variable, assuming that "taxes are more likely to be the factor which is adjusted" (Blanchard, 1990: p. 13).\textsuperscript{23}

In theory, the necessary adjustments can be achieved through a cut in public spending, an increase in taxes or resorting to monetary financing. In practice, the empirical studies usually employ the tax rate as the endogenous variable, assuming that "taxes are more likely to be the factor which is adjusted" (Blanchard, 1990: p. 13).\textsuperscript{23}

Define $t^*$ as the constant average tax rate (i.e., in practical terms, the ratio of all non-monetary government revenues to GDP) necessary to stabilise the ratio of debt to GDP during a certain time period. The difference between this 'sustainable' tax rate and the actual implicit tax rate $t$ indicates the amount of the necessary adjustment, at each period, providing a measure of 'fiscal pressure'. This 'tax gap' may be derived directly from equation (2.13),

$$
d_t - E_t \delta_{t,n} d_s^{t,n} - E_t \sum_{j=1}^{n} \delta_{t,j} \left[ s_{t+j} + \Delta m_{t+j} + m_{t+j-1} \left( 1 - \frac{1}{(1 + \pi_{t+j})(1 + \psi_{t+j})} \right) \right]
$$

(2.18) $t^* - t = \frac{E_t \sum_{j=1}^{n} \delta_{t,j}}{E_t \sum_{j=1}^{n} \delta_{t,j}}$

This indicator can be calculated for an infinite set of different time horizons, from the more 'myopic' one-year indicator until the infinite horizon case. Evidently, the longer the time-horizon, the higher the probability of errors, since more assumptions have to be made about the future path of the variables. Blanchard \textit{et al.} (1990) suggest and use three different indicators for, respectively, one, five and forty years.\textsuperscript{24} This last, longer period indicator allows them, for instance, to consider the fiscal consequences arising from the demographic pressures of an ageing population. OECD (1990) and Uctum and Wickens

countries with lower GDP per capita.

\textsuperscript{22} The OECD has been particularly interested in finding an appropriate and operational indicator of fiscal policy sustainability, based on projections of the relevant fiscal variables. Apart from the ones examined here, see for instance the 'simulations of public debt' illustrated in Chouraqui \textit{et al.} (1986), or the indicator corrected for the influence of the business cycle, described in Blanchard \textit{et al.} (1990). Because they are presently not widely known and applied, these indicators will not be further explored here. Besides, they are in essence not very different from the ones presented in this section. Buiter (1995b: p. 1) claims that all these proposed indicators are fundamentally an attempt to correct the "myopic (...) traditional financial performance criteria" used under the IMF programs.

\textsuperscript{23} Yet, Bohn (1991a) finds evidence that in the United States, a budget deficit is, on average, eliminated in 50-65\% by a decrease in spending and in 35-50\% by an increase in taxes.

\textsuperscript{24} Blanchard (1990) also defines a 'primary gap', requiring only current values, which is exactly the same as the one defined in section 2.3, the difference between the left and the right-hand side of equation (2.12).
(1996), for example, have chosen the three and five years horizon respectively, a choice determined by the availability, at the time of publication, of the OECD projections for the relevant variables.

Artis and Marcellino (1998) test for the accuracy of these OECD estimates. They found that they are rather accurate for most countries, although surpassed by a simple driftless random walk model. Buiter et al. (1993) compare the official forecasts from the OECD and the EC, and note that they can be significantly different in some countries. Therefore, instead of using official forecasts, another possible approach is to construct time series ARMA models for the variables of interest, using them to forecast future values (see Chouraqui et al., 1986, and Artis and Marcellino, 1998, for example). However, this procedure does not take into account expected changes in policy, the main publicised advantage of the sustainability tests examined in this section.

Whatever the source of the forecasts, a positive value of the measure \((t'-t)\) is an indication of fiscal unsustainability, suggesting the need for an increase in primary surpluses through an increase in taxes or a decrease in public expenditures, an increase in monetary financing, or even, in the limit, debt repudiation.

The relative importance of the tax 'gap' depends on each country's actual fiscal situation. A positive gap will be much more difficult to fill in a country with already high tax rates and strong commitments in terms of public expenditures. In effect, if the amount of the required fiscal adjustment is too high, the government may not be able to carry the necessary correction in taxes or public expenditures, forcing it to resort to monetization or repudiation. Since the risk of resorting to one of these more 'drastic' solutions depends on the fiscal burden already imposed on the economic agents, one possible and simple 'standardised' indicator of this relative risk is the measure \((t'-t)/(1-t)\), where the denominator represents the theoretical amount of resources which the government may still confiscate.

The same approach could be followed to find the 'sustainable spending rate' or even the 'sustainable' growth rate of base money. This is why some researchers (Buiter et al., 1993, and Roubini, 1995, for example) prefer to use as indicator the so called 'primary gap', defined as the difference between the constant 'sustainable primary balance' and the actual primary balance, in terms of GDP. This indicator recommends the size of the minimum adjustment required to avoid a rising rate of debt-GDP.
The empirical results from the tests employing this measure of sustainability with a collateral constraint are not consensual. This is a natural reflection of the very strong assumptions about the future behaviour of the fiscal variables, on which they are based upon. Blanchard et al. (1990), Fase and Wellink (1990), and OECD (1990), found that most OECD countries have sustainable fiscal policies in the medium term. The consistent exceptions are Italy, Greece and the Netherlands. Contrarily, and following the same methodology but with more recent data, Roubini (1995) concludes that most OECD countries are presently on an unsustainable path.

For the European Union’s members specifically, Buiter et al. (1993) found problems in Greece, Italy and Germany, while Uctum and Wickens (1996) conclude that, with the exception of Spain, all countries have sustainable fiscal policies. A similar study, presented in European Commission (1994), exclusively for the case of Greece, found problems of sustainability in the medium term, beginning in 1980.

As already mentioned above, the main problem with this indicator of sustainable fiscal policies is the availability and quality of the projections used to compute them. These projections are widely available for a short-term horizon of one to five years. However, as reported for instance in Blanchard et al. (1990) and OECD (1990), the conclusions from the tests using these short-term horizons are not significantly different from the ones obtained with solutions for the steady-state explored in section 2.3.1. For longer time horizons, the availability and credibility of the projections are highly uncertain.

2.4. Sustainability with a transversality condition

The methods examined in the previous section test the double constraint of intertemporal balance and stability of the debt-GDP ratio, in a restrictive view of sustainability. A more general approach, adopted in the majority of the literature, is to test the IBC in an infinite horizon, based on the long-term properties of its fundamental variables. The debt ratio

25 The motivation for this approach may have been given in October, 1985 by R. Lucas' (quoted in Hansen and Sargent, 1991: p. 9) challenging question of "what restrictions would be imposed on a joint stochastic
is not required to stabilise, but its growth rate should not exceed the interest rate. This
approach was initiated by Hamilton and Flavin (1986) and has developed into distinct
testing procedures, which will be systematically analysed in this section.

Since the sustainability condition is interpreted as a long run relationship, it suggests the
use of methodologies derived from the unit root and cointegration literature to test it
empirically. A set of tests can be employed to examine the long run implications of the
current fiscal and financial policies of the government. The econometric techniques used
have followed closely the continuous advances in time series econometrics. In fact, most
of the empirical work in this research area has been driven by continuous but small
methodological changes.

Basically, the empirical studies investigate the order of integration and cointegration of
the series of debt and deficit, or the various components of the latter - government
expenditures, tax revenues, and seigniorage. The main objective is to detect whether a
fiscal policy is sustainable if the fiscal variables continue to follow the stochastic
processes estimated in the econometric tests. As already mentioned, the tests of the IBC
are not tests of government solvency, but rather of sustainability of its fiscal policy. To
test for solvency one would have to consider all possible future values of the relevant
fiscal variables, and check if there was any possibility of not satisfying the IBC.

The seigniorage revenues are usually not explicitly considered in these sustainability
tests. Most authors do not differentiate between tax and seigniorage revenues, others
simply disregard them as insignificant. This usual procedure will also be adopted in the
following sections, with $NS_t$ denoting the net-of-interest surplus and $TT_t$ the total
government revenues. Chapter 5 will examine the consequences of differentiating both
sources of revenue, which clearly have different effects on the economy.

Several different criteria could have been used to classify the long-term testing
procedures examined in this section. Among others, the tests can be divided according to
a chronological order, testing conditions, econometric techniques, variables used, or
assumptions on the interest rate. This last criterion was chosen here, since it seems the
most important factor of differentiation between the testing approaches.

---

process describing the net-of-interest government expenditures and taxes by the assumption of present
value budget balance"
In section 2.4.1, a constant interest rate is assumed. This is a common characteristic of most initial testing approaches, but it may lead to spurious results given the implausibility of this strong assumption. Three main testing procedures are analysed in this section. The first two connect sustainability of fiscal policies with the stationarity of, respectively, the series of net-of-interest and total deficit. The third is a more extensive approach, due to Hansen, Roberds and Sargent (1991). It tests a group of cross-equation restrictions on the VAR representation of the debt and net-of-interest surplus.

The following section, 2.4.2, examines the alternative tests for when the assumption of a constant interest rate is dropped. The first solution was presented by Wilcox (1989), who basically avoids this question by using the actual values of the interest rate on public bonds to discount the debt series back to a certain base period, and testing this discounted series. The second solution, in 2.4.2.2, consists of testing the stationarity of the first-differenced debt series. It is shown that this test is robust to a variable interest rate specification.

All the above approaches to test the IBC rest on the assumption of a nonstochastic economy. Section 2.4.3 presents the stochastic version of the model, which differs from the nonstochastic case mainly by the treatment of the discount factor. The first main implication is that the future flow of budget surpluses cannot be discounted by the 'risk-neutral' interest rate on the government bonds. Bohn (1995) claims that abandoning the unrealistic assumption of a deterministic economy invalidates all the above mentioned tests of sustainability. However, it is demonstrated that, under certain assumptions, some of these tests are still valid in a stochastic economy.

2.4.1. The assumption of a constant interest rate

All the methodologies reviewed in this section assume a constant interest rate when deriving the testing conditions of sustainability of fiscal policies. These distinct conditions provide a secondary rule for classifying the tests into the following three groups, which also obey a chronological order.
2.4.1.1. The first tests: stationarity of the net-of-interest deficit

The most straightforward evaluation of the null hypothesis that equations (2.6) and (2.7) hold, is to test the order of integration of the series of debt $D_t$ and net-of-interest surplus $NS_t$. If the government does not engage in bubble finance, i.e., if equation (2.6) is satisfied, than (2.7) implies that $D_t$ and $NS_t$ must be integrated of the same order. This assumes the notion of "stationarity of the undiscounted surplus being sufficient for stationarity of the sum of expected discounted surpluses, assuming that the real interest rate is positive" (Wilcox, 1989: p. 297).

This was essentially the test conducted in the seminal paper of Hamilton and Flavin (1986), where they introduced the use of stationarity tests to analyse long-run limitations on debt financing. The alternative to the hypothesis of sustainability is to consider the class of debt processes satisfying

$$E_t \lim_{n \to \infty} \frac{D_n}{(1+r)^n} = A_0,$$

where $A_0$ can be any positive constant. Substituting in (2.5), with a constant discount factor, yields

$$D_t = A_0(1+r)^t + \sum_{j=1}^{n} E_t \frac{NS_{t+j}}{(1+r)^j}.$$  

To satisfy the transversality condition and the IBC, $A_0$ must be zero. If the last term in (2.20) is stationary, $A_0$ will be zero only if $D_t$ is also stationary. Therefore, Hamilton and Flavin's testing procedure amounts simply to a unit root test on both the net-of-interest

26 The tests of the government IBC were inspired in the literature on asset price bubbles. In the asset pricing theory, the transversality condition states that the present discounted value of the asset's terminal price goes to zero in the limit. As shown below, Campbell and Shiller's (1987) tests of a present value model for bonds and stocks, for example, are very similar to tests of the IBC. A bubble in the price of an asset can be defined as the difference between the price of the asset and its fundamental value, given by the present discounted value of the expected flow of returns from that asset (Blanchard and Fisher, 1989). In equation (2.5), for example, the first term on the right-hand side is the bubble component, whereas the last term represent the fundamental element. In this sense, bubbles can be compared to Ponzi games (see, for example, O'Connel and Zeldes, 1988).

27 Unlike almost all posterior authors, Hamilton and Flavin regarded their tests as a means of verifying whether the government must satisfy the IBC. Their interpretation was that a violation of this constraint would not have indicated unsustainable fiscal policies and a necessity to correct them, but simply that this condition does not have to be satisfied.
surplus and the stock of debt. If both time series are stationary, equation (2.6) holds and the intertemporal budget constraint is satisfied.

The authors applied the test to the United States’s fiscal stance during the period 1960-84, and rejected the unit root hypothesis for the series of debt and deficit, concluding in favour of the sustainability of the debt process. However, the power of their tests is low, given the small number of observations. Besides, their rejection of the null hypothesis of nonstationarity was due to the choice of a 10 per cent significance level. At the more usual 5 per cent level, the test indicates nonstationarity in both series. Moreover, Kremers (1988) have shown that the results from the Hamilton and Flavin’s ADF test are not satisfactory, due to a wrong specification of the lag length employed to eliminate serial correlation in the residuals. With a more correct specification of the test, and using the same data set, he found evidence of unsustainability.

Soon after the leading paper of Hamilton and Flavin, another test was proposed by Hakkio and Rush (1986), based on the direct analysis of equation (2.1), reproduced below, and using several tests for cointegration suggested by Engle and Granger (1987)

\[
G_t - \left( T_t + \frac{\Delta M^N_t}{P_t} \right) = D_t - (1 + r_t) D_{t-1}
\]

If the stochastic processes \(G_t\) and \(TT_t\) are nonstationary, they must be cointegrated so that the transversality condition holds. Intuitively, if the public expenditure in a certain moment increases permanently by a certain value, public revenues must also increase accordingly, in order to keep debt from exploding. If they are not cointegrated, their difference, equal to \(D_t - (1 + r_t) D_{t-1}\), will be nonstationary.

However, one major problem with this test is that cointegration between the two time series is just a necessary, but not sufficient, condition for sustainability. It is also necessary that the cointegration factor equals one to ensure that public expenditures are not growing at a higher rate than revenues. Otherwise, if the drift term or trend in \(G_t\) is larger than that in \(TT_t\), the right hand side of equation (2.21) will also have a positive drift

---

18 Applying this same test procedure to six European Union countries, Heinemann (1993) found evidence of unsustainability, surprisingly, only in Germany and France.
or trend. This means that, on average, \( D_t > (1+r_t) D_{t+1} \), which indicates that the government is involved in Ponzi schemes, paying all the maturing debt by further borrowing. In other words, the growth rate of the debt is higher than the real interest rate on that debt, violating the transversality condition.

Like in Hamilton and Flavin (1986), the tests performed by Hakkio and Rush (1986) with post-war US data have also supported the hypothesis of sustainability. Essentially, these two initial papers suggest that stationarity of the series of debt and surplus, including seigniorage, is a necessary condition to meet the IBC. As will be shown below, later analyses have established that the series of debt and net-of-interest surplus do not have to be stationary, provided they are both I(1) and cointegrated.

### 2.4.1.2. Stationarity of the total deficit

A different line of research was initiated with the papers of Trehan and Walsh (1988, 1991), where they show that stationarity of the undiscounted series of debt and net-of-interest surplus, although sufficient, is not a necessary condition to guarantee intertemporal budget balance. Their test is less restrictive, allowing for nonstationarity of the series of debt and seigniorage-augmented primary surplus, provided they are cointegrated and the total surplus (including interest payments) is stationary.

The logic behind this assertion is intuitive, as demonstrated for example by Vanhorebeek and Van Rompuy (1995). Suppose that the net-of-interest surplus \( NS_t \) is affected by a significant negative temporary shock (either on primary expenditures \( G_t \), tax revenues \( T_t \) or seigniorage \( \Delta M_t^N/P_t \)), but remains a stationary process. This negative shock immediately affects the stock of debt, setting up a self-fulfilling cycle of additional interest expenditures \( r_tD_{t+1} \) and increasing debt. The series of public debt will therefore be a nonstationary process. To comply with the transversality condition (2.6), the government now needs to present sufficient primary surpluses, including seigniorage, to

---

29 Caporale (1993a) has also performed these tests of cointegration between primary expenditures and revenues, but based on wrong conclusions taken from equation (2.29) below.

30 If they are stationary, it will have implications for Barro’s tax-smoothing hypothesis (Barro, 1979). Trehan and Walsh (1988) have shown that the deficit net of interest payments will generally be nonstationary under Barro’s hypothesis.
pay those additional interest expenditures. For this to happen, it is necessary that the total deficit, \( r_D t, t \sim NS_t \), is a stationary process.

Starting from the definition of the stochastic process followed by the net-of-interest surplus as a ratio of GDP, and supposing that the IBC holds, it is possible to derive testable hypothesis for the joint process followed by the series of debt and net-of-interest surplus, or by the various components of the total surplus. The basic reasoning is that if the net-of-interest surplus and the stock of debt are in fact linked by the IBC, then if the process \( ns_t \) follows a trend, the process \( d_t \) must have the same trend. Although they may fluctuate independently in the short run, the IBC demands that they exhibit a common trend in the long run. For example, suppose that the surplus follows the deterministic linear function

\[
ns_t = \alpha + \beta t.
\]

Substituting in (2.11) \( d_t = \sum_{j=1}^{\infty} E_t, \delta, j, ns_{t+j} \), and considering a constant and positive, growth-rate-corrected interest rate \( \rho \), the discount factor becomes \( \delta_j = (1+\rho)^j \), producing

\[
d_t = \frac{\beta(1+\rho)}{\rho^2} + \frac{\alpha}{\rho} + \frac{\beta}{\rho} = \alpha \rho + \beta(1+\rho) + \frac{\beta}{\rho} t.
\]

This demonstrates that the series of debt must follow the same trend of the net-of-interest surplus, scaled by the growth-adjusted interest rate.

In a more general case, let \( ns_t \) be characterised by the following ARIMA process:

\[
(1-\phi L) ns_t = \theta(L) \varepsilon_t,
\]

where \( 0 < \phi < (1+\rho) \), \( L \) is the lag operator, and \( \theta(L) = \sum_{j=0}^{\infty} \theta_j L^j \), assumed to be square summable \( \sum_{j=0}^{\infty} \theta_j^2 < \infty \).\(^{31}\) The right-hand side of (2.24) is a zero mean stationary stochastic process, represented as a moving average of white noises.\(^{32}\) This specification of the process followed by the

---

\(^{31}\) In reality, the condition should not allow the possibility of \( \phi > 1 \), since in this case the net-of-interest surplus will not be stationary even after first differencing.

\(^{32}\) Trehan and Walsh (1988) allowed for a deterministic trend in the surplus stochastic process \([1-\phi L] ns_t = \mu + \theta(L) \varepsilon_t\). This implies the additional necessary condition that the initial net-of-interest surplus must be sufficient to cover the interest payments on the initial stock of debt.
seigniorage-augmented primary surplus allows this variable to be stationary, if \(0 \leq \phi < 1\), or nonstationary, in the case where \(1 \leq \phi < (1+\rho)\).

Substituting again in (2.11) yields

\[
(2.25) \quad d_t = E_t \left[ \frac{\phi}{(1+\rho)} n s_t + \frac{\theta(L)}{(1+\rho)} e_{t+1} + \frac{\phi}{(1+\rho)^2} n s_t + \frac{\theta(L)}{(1+\rho)^2} e_{t+1} + \frac{\theta(L)}{(1+\rho)^2} e_{t+2} + \ldots \right],
\]

which after some simple algebraic manipulation can be reduced to

\[
(2.26) \quad d_t = \frac{\phi}{(1+\rho)} n s_t + \frac{1}{(1+\rho)} E_j \sum (1+\rho)^{-j} L^{-j} \sum \theta(L) e_{t+1}.
\]

Using the formula derived in Hansen and Sargent (1980) to calculate the expected present value of a stationary moving average process, which shows that

\[
(2.27) \quad \sum \lambda^k E_t \chi(L) e_{t+k} = \frac{\chi(L) - \lambda \chi(L) L^{-1}}{1 - \lambda L^{-1}} e_t,
\]

where \(e_t\) is white noise and \(\chi(L) = 1 + \chi_1 L + \chi_2 L^2 + \ldots\), and square summable, the following expression is obtained

\[
(2.28) \quad d_t = \frac{\phi}{(1+\rho)} n s_t + \frac{1}{(1+\rho)} \left[ \frac{\theta(L) - (1+\rho)^{-1} \theta((1+\rho)^{-1}) L^{-1}}{1 - (1+\rho)^{-1} L^{-1}} \right] e_{t+1}.
\]

Using equation (2.24), describing the stochastic process followed by the net-of-interest surplus, in equation (2.28), will finally yield

\[
(2.29) \quad d_t = \frac{1}{(1+\rho)} n s_{t+1} + A(L) e_{t+1},
\]

where

\[
A(L) = \frac{1}{(1+\rho)} \left[ \frac{\theta(L) - (1+\rho)^{-1} \theta((1+\rho)^{-1}) L^{-1}}{1 - (1+\rho)^{-1} L^{-1}} - \theta(L) \right].
\]

If \(n s_t\) is a stationary time series process, then the IBC holds only if the process \(d_t\) is also stationary. This is precisely the situation tested in Hamilton and Flavin (1986). Otherwise, if \(n s_t\) is nonstationary, equation (2.29) shows that if the surplus follows the process described in (2.24), if the conditional expectation of the interest rate is constant, and if the intertemporal budget constraint (2.11) is satisfied, then \(d_t\) must also be

---

33 See their equation (6) and the appendix A or, alternatively, Sargent (1987: p. 304).
nonstationary, and \( d_t \) and \( ns_{t+1} \) are cointegrated. When \( \phi \) equals unity, the cointegrating term will be the inverse of the interest rate.

Conversely, it can also be demonstrated that the inverse necessarily holds, i.e., that if the series of debt and net-of-interest surplus are cointegrated, with a cointegrating parameter \( \alpha \), then the transversality condition is satisfied (Trehan and Walsh, 1991). Assume the following linear combination of the variables \( d_{t-1} \) and \( ns_t \)

\[
d_{t-1} - \alpha ns_t = B(L)e_t,
\]

where \( B(L)e_t \) is a stationary process. Multiplying both sides of (2.30) by \( (1-\phi L) \) yields

\[
(1-\phi L) d_{t-1} - \alpha (1-\phi L) ns_t = (1-\phi L)B(L)e_t.
\]

Substituting (2.24) in (2.31) and rearranging produces the expression

\[
(1-\phi L) d_{t-1} = \alpha \theta(L) e_t + (1-\phi L)B(L)e_t = C(L)e_t,
\]

where \( C(L) = \alpha \theta(L) + (1-\phi L)B(L) \). Solving (2.32) forwards gives

\[
d_{t+n} = \phi^n d_t + \sum_{j=0}^{n-1} \phi^j C(L)e_{t+n+1-j}.
\]

Finally, substituting in the transversality condition (2.10), yields

\[
\lim_{n \to \infty} \left[ \frac{\phi}{(1+\rho)} \right] d_t + \lim_{n \to \infty} \sum_{j=0}^{n-1} \frac{\phi^j}{(1+\rho)^j} C(L)e_{t+n+1-j}.
\]

Since it was initially assumed in (2.24) that \( 0 \leq \phi < (1+\rho) \), these limits will always equal zero.

In sum, if the net-of-interest surplus is a stationary process \( 0 \leq \phi < 1 \), the debt process must also be a stationary process, in order to maintain the intertemporal solvency condition. If the surplus is a difference stationary process, then \( d_t \) and \( ns_{t+1} \) must be cointegrated. In this particular case with \( \phi = 1 \), the cointegrating vector will be \( [1 \ -1/\rho] \). Equivalently, the condition is that the stock of debt and the net-of-interest surplus should be cointegrated, with cointegrating vector \( [-\rho \ 1] \). These were essentially the tests carried out by, among others, MacDonald and Speight (1990) for the UK, Smith and Zin (1991) for Canada, Trehan and Walsh (1991) and MacDonald (1992) for the US, Baglioni and Cherubini (1993) for Italy, and McNellis and Siddiqui (1993) for New Zealand.
The analysis of a cointegrating relationship between the series of debt and seigniorage-augmented primary surplus is a possible way of testing sustainability of fiscal policies. However, equation (2.29) allows further interpretations and different ways to assess the sustainability hypothesis. For example, rearranging the first two terms in (2.29) and considering the surplus to be a difference stationary process yields

$$d_t - \frac{1}{\rho} n_{s_{t+1}} = \frac{1}{\rho} (\rho d_t - ns_{t+1}) = A(L) e_{t+1}.$$  \hspace{1cm} (2.35)

From this expression a second possible testing procedure can be derived indirectly from equation (2.29). To ensure sustainability, the government's total deficit (the expression in parentheses) must be stationary. Trehan and Walsh (1988) used these two theoretically equivalent tests, but had some problems to obtain convincing conclusions, since the results from the cointegration test, in the first case, and the unit root test, in the second case, were not consistent.

Equivalently, it is possible to derive a third type of testable implications from (2.29). Decomposing the deficit in its main components yields

$$\frac{1}{\rho} (\rho d_t - n_{s_{t+1}}) = \frac{1}{\rho} \left( \rho d_t + g_{t+1} - t_{t+1} - \frac{\Delta \hat{M}^n}{GDP} \right).$$  \hspace{1cm} (2.36)

The condition of stationarity of the total deficit requires that the series of debt, government primary expenditures and government total revenues are cointegrated, with cointegrating vector $[\rho 1 -1]$. Intuitively, if the stock of debt rises in a certain period, the additional interest payments must be matched either by a fall on public primary expenditures or by a rise in revenues, in order to satisfy the IBC. Hénin and Garcia (1996) applied this test to a group of six OECD countries and found evidence of sustainable policies only in Japan.

Disregarding the revenues from money creation in equation (2.36), Bohn (1991a) concluded that stationarity of the first difference of the series of debt ($\Delta D_t$) requires cointegration between the series of tax revenues $T_t$, non-interest expenditures $G_t$ and debt $D_t$, with a cointegrating vector $[1 -1 -\rho]$. On the other hand, isolating the revenues from base money creation, Trehan and Walsh (1988) also tested whether the components of the interest inclusive budget deficit $\rho d_{t-1} + g_t$, $t_t$ and $m_t - m_{t-1}/(1 + \eta)(1 + \psi)$ are cointegrated, with cointegrating vector $[1 -1 -1]$. 

36
Instead, and this is a fourth alternative testable hypothesis, the sustainability condition may be that total government expenditures $(G_t+r_tD_t)$ and total revenues $(T_t+\Delta M_t/P_t)$ are cointegrated, with cointegrating vector $[1\ -1]$. This is equivalent to the condition proposed and tested by Hakkio and Rush (1991a) in a different context, which will be analysed with more detail in section 2.4.2.2, below.

2.4.1.3. The Hansen, Roberds and Sargent test

The test proposed by Hansen et at. (1991) is based on the analysis of the stochastic properties of the series of debt and net-of-interest surplus, i.e., including seigniorage but excluding interest payments. For sustainability to hold, two different alternative groups of restrictions must be verified, depending on whether those two variables are both stationary or nonstationary.

Their test is also analytically derived from the IBC, either equation (2.7) or (2.11). For reasons of algebraic simplification, the demonstration of this test follows from equation (2.7), where the whole expression in parenthesis, the net-of-interest surplus is consolidated in the term $NS_t$, yielding

$$D_t = \sum_{j=0}^{\infty} E_t \delta_{t+j+1} NS_{t+j+1}.$$  

As noted before, the current value of debt must equal the discounted sum of all expected future net-of-interest surpluses. These expectations are conditioned on the full public information set available, which consists of past observations not only of the surplus but also of the debt series. This point, central to the approach of Hansen et al. (1991), will be resumed later in this section.

If the series of debt and net-of-interest surplus are stationary, equation (2.37) is the basis of the test. If they are non-stationary, this equation must be adjusted. Multiplying both terms by the constant real interest rate and subtracting the current net-of-interest surplus yields

$$rD_t - NS_t = r \sum_{j=0}^{\infty} E_t \delta_{t+j+1} NS_{t+j+1} - NS_t.$$  

37
or, with a slight but helpful modification, feasible because of the assumed constancy of the discount factor,

\[(2.39) \quad rD_t - NS_t = r\delta_t \sum_{j=0}^{\infty} E_t \delta_{t+j} NS_{t+j+1} - NS_t,\]

which is equivalent to

\[(2.40) \quad rD_t - NS_t = \sum_{j=0}^{\infty} E_t \delta_{t+j} NS_{t+j+1} - \delta_t \sum_{j=0}^{\infty} E_t \delta_{t+j} NS_{t+j+1} - NS_t,\]

since \(r\delta_t = (1 - \delta_t).\) Aggregating the last two terms, the following expression is obtained,

\[(2.41) \quad rD_t - NS_t = \sum_{j=0}^{\infty} E_t \delta_{t+j} NS_{t+j+1} - \sum_{j=-1}^{\infty} E_t \delta_{t+j} NS_{t+j+1},\]

which finally simplifies to

\[(2.42) \quad rD_t - NS_t = \sum_{j=0}^{\infty} E_t \delta_{t+j} \Delta NS_{t+j+1} .\]

This expression states that the current value of the total deficit must be exactly matched by the discounted sum of all expected changes in the net-of-interest surplus. This is an alternative way of interpreting the IBC.\(^{34}\)

If the net-of-interest surplus is a stationary process, equation (2.37) shows that the debt series must also be stationary.\(^ {35}\) Equivalently, if the surplus is I(1), and therefore its first difference is stationary, equation (2.42) shows that the stochastic process \((rD_t - NS_t)\) must also be stationary. In this case, \(D_t\) and \(NS_t\) must be cointegrated.

Whatever the univariate stochastic properties of the series of debt and net-of-interest surplus, Campbell and Shiller (1987) and Hansen et al. (1991) have shown that equations (2.37) and (2.42) impose certain restrictions on the moving average representation of the joint stochastic processes \(\{D_n, NS_t\}\) or \(\{rD_t - NS_t, \Delta NS_t\}\), depending respectively on

----

\(^{34}\) Equivalent to the condition proposed by Campbell (1987) for the consumer obeying the permanent income theory. Today's savings must equal the discounted sum of all expected future changes (reductions) in income.

\(^{35}\) It can be shown that the infinite weighted sum (with the weights decreasing geometrically) of expected conditioned values of stationary processes is also a stationary process (see section 2.4.1.2).
whether the net-of-interest surplus is stationary or not. In what follows, the second hypothesis is considered, but the methodology is basically equivalent in both cases.

These equations also impose restrictions on the corresponding VAR representation (which results from the inverted moving average representation) of the joint process $y$, depicted by the bivariate system

\begin{equation}
\begin{bmatrix}
D_t - NS_t \\
D_{t-1} - NS_{t-1} \\
\vdots \\
D_{t-p+1} - NS_{t-p+1} \\
\Delta NS_t \\
\Delta NS_{t-1} \\
\vdots \\
\Delta NS_{t-p+1}
\end{bmatrix} = \begin{bmatrix}
a_1 & a_2 & \cdots & a_p & b_1 & b_2 & \cdots & b_p \\
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
c_1 & c_2 & \cdots & c_p & d_1 & d_2 & \cdots & d_p \\
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\begin{bmatrix}
D_{t-1} - NS_{t-1} \\
D_{t-2} - NS_{t-2} \\
\vdots \\
D_{t-p} - NS_{t-p} \\
\Delta NS_{t-1} \\
\Delta NS_{t-2} \\
\vdots \\
\Delta NS_{t-p} \\
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0
\end{bmatrix} + \begin{bmatrix}
u_{t} \\
\vdots \\
0 \\
\vdots \\
0
\end{bmatrix},
\end{equation}

where $a(L)$, $b(L)$, $c(L)$ and $d(L)$ are polynomials in the lag operator $L$, with a maximum lag order of $p$ (and removing the mean of each element of $y$). For convenience, this can be stacked to form the system

\begin{equation}
\begin{bmatrix}
rD_t - NS_t \\
rD_{t-1} - NS_{t-1} \\
\vdots \\
rD_{t-p+1} - NS_{t-p+1} \\
\Delta NS_t \\
\Delta NS_{t-1} \\
\vdots \\
\Delta NS_{t-p+1}
\end{bmatrix} = \begin{bmatrix}
a_1 & a_2 & \cdots & a_p & b_1 & b_2 & \cdots & b_p \\
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
c_1 & c_2 & \cdots & c_p & d_1 & d_2 & \cdots & d_p \\
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\begin{bmatrix}
rD_{t-1} - NS_{t-1} \\
rD_{t-2} - NS_{t-2} \\
\vdots \\
rD_{t-p} - NS_{t-p} \\
\Delta NS_{t-1} \\
\Delta NS_{t-2} \\
\vdots \\
\Delta NS_{t-p} \\
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0
\end{bmatrix} + \begin{bmatrix}
v_{t} \\
\vdots \\
0 \\
\vdots \\
0
\end{bmatrix},
\end{equation}

abbreviated by

\begin{equation}
z_t = Az_{t-1} + v_t,
\end{equation}

where $A$ is the 'companion matrix' of coefficients.

The expected value of $z_{t+j}$, conditional on the information available at time $t$, is a linear projection of the elements of \{ $rD_{t-1} - NS_{t-1}$, $\Delta NS_{t-1}$, $rD_{t-2} - NS_{t-2}$, $\Delta NS_{t-2}$, ... \}, given by

\begin{equation}
E_t z_{t+j} = A^{j+1} z_t.
\end{equation}

This is the standard forecasting formula for the variables in the VAR. Define two $2p \times 1$ vectors, $f$ and $h$, with all elements zero except, respectively, the first element in $f$ and the $p+1$th in $h$, which equal unity. From equation (2.42),

\begin{equation}
f^t z_t = rD_t - NS_t = \sum_{j=0}^{p} E_t \delta_{t,j} \Delta NS_{t+j+1},
\end{equation}

39
and from equations (2.45) and (2.46),

\[(2.48) \quad h'z_t = \Delta NS_t \Rightarrow h' A^{j+1} z_t = E_r \Delta NS_{t+j+1}.\]

Substituting (2.48) in (2.47) yields

\[(2.49) \quad f'z_t = \sum_{j=0}^{\infty} \delta_j h' A^{j+1} z_t.\]

This expression merely attests the required equality between \(rD_t-NS_t\) and the unrestricted forecast of all the future discounted \(\Delta NS_t\) from the VAR, computed by applying the intertemporal prediction formula (2.46). Using the expression of the infinite sum of a geometric progression, this is equivalent to

\[(2.50) \quad f'z_t = h' A (1 - \delta_{t,1} A)^{-1} z_t \]

or, in linear form, this requires that for any \(z_t\),

\[(2.51) \quad f'(1 - \delta_{t,1} A) = h' A,\]

From (2.44), the following cross-equation restrictions can be derived:

\[(2.52) \quad 1 - \delta_{t,i} a_i = c_i, \text{ and} \]

\[\delta_{t,1} a_i = c_i, \quad i = 2, 3, \ldots, p, \text{ and} \]

\[\delta_{t,1} b_i = d_i, \quad i = 1, 2, \ldots, p.\]

Rejection of the restrictions in (2.52) indicates that the expectations of all future changes in the net-of-interest surplus are not sufficient to justify the current level of \(rD_t-NS_t\). Campbell and Shiller (1987) proposed and empirically tested these restrictions in the more general class of the 'present value models'

\[Y_t = \theta (1 - \delta) \sum_{i=0}^{\infty} E_r \delta^i x_{t+i} + c.\]

They used this model to study the relation between long-term and short-term interest rates, and the relation between the stock prices and dividends. Campbell (1987) uses it to examine the permanent income theory of consumption.
The IBC model is a particular version of these present value models, and it was presented by Hansen et al. (1991) and tested by MacDonald and Speight (1990) and Roberds (1991). MacDonald and Speight applied this testing procedure to the United Kingdom, using quarterly data for the period 1961-86, and strongly rejected the restrictions in (2.52). Roberds found that the government’s IBC is also not satisfied in the United States, in the period 1948-86 with quarterly data. This stands in sharp contrast with the previous findings of Hamilton and Flavin (1986), Hakkio and Rush (1986) and Trehan and Walsh (1988), also for the US.

However, before testing these cross-equation coefficient restrictions, it is convenient to start by testing another implication of the analysis above, which is that $rD_t - NS_t$ must Granger-cause $\Delta NS_t$. Campbell and Shiller (1987) provide an intuitive explanation of this condition. The left-hand side of equation (2.42) is the optimal forecast of the weighted sum of all expected first-differenced net-of-interest surpluses. These forecasts will be superior if the values of the current and past values of $rD_t - NS_t$ are included in the information sets than if these include only current and past values of the $MS_t$ variable itself. In other words, $rD_t - NS_t$ Granger-causes $\Delta NS_t$ because the agents have superior information than that contained in lagged $\Delta NS_t$.

If however Granger non-causality cannot be rejected, Roberds (1991) has shown for the particular case of the IBC that the cross-equation restrictions (2.51) do not make sense, because of stochastic singularity of $\{rD_t - NS_t, \Delta NS_t\}$, and one is back to the single-equation tests presented earlier by Hamilton and Flavin (1986), for example.

2.4.2. A variable interest rate

All the studies presented above start by imposing the assumption of a constant interest rate, throughout the whole period analysed, in order to make equation (2.3) more tractable. This extremely simplifying assumption may affect significantly the results of

---

36 Hansen et al. (1991) claim that if data on the stock of government debt is not available, the IBC is theoretically not testable.

37 MacDonald and Speight (1990), for instance, found that $rD_t - NS_t$ does not appear to Granger-cause $\Delta NS_t$, which was seen as a support for their rejection of the cross-equation restrictions.
the empirical tests of sustainability. In fact, some papers report for instance a negative point estimate of the real interest rate in the cointegration regression (Haug, 1991, and Smith and Zin, 1991, for example), when the interest rate is one of the elements of the cointegrating vector to be estimated. This finding casts some doubts on the validity of those models, since it is inconsistent with the hypothesis of intertemporal budget balance of equations (2.6) and (2.7), which implies a positive expected real interest rate.

This section introduces some more general tests allowing, more realistically, the presence of variable interest rates. The first testing approach manages to avoid the need for a particular specification of the interest rate by using the actual values of the rate on public bonds to compute a discounted debt series. In 2.4.2.2 an alternative method of deriving the testing conditions is presented, where the restriction of a constant interest rate may also be relaxed.

2.4.2.1. Stationarity of the discounted debt (Wilcox's test)

One process of avoiding the assumptions on the expected interest rates is to use ex post realized interest rates to construct a discounted debt series. In this sense, Wilcox (1989) was the first to allow for stochastic variations in the real interest rate, by measuring ex post holding rates of return on the debt and then discounting the series of debt back to a fixed reference date. Multiplying (2.1) by the discount factor

\[ \delta_{t+i} = \prod_{i=0}^{t-1} \left( \frac{1}{1+r_i} \right), \]

and expressing \( D'_i \) and \( NS'_i \), respectively as the discounted values of the debt and net-of-interest surplus, yields

\[ D'_i = D'_{i+1} - NS'_i, \]

which is basically equivalent to the identity in (2.1) with the only difference that the variables of period \( t \) are now discounted back to period 0. By recursive substitution, as in (2.3), this becomes the intertemporal identity

---

38 It implies, for example, that the interest rate is not sensitive to the amounts of deficit and debt, a question empirically examined in chapter 6.
The intertemporal budget constraint holds for an infinite horizon if the expectation of the discounted debt approaches zero as time approaches infinity, i.e., in equation form,

\[(2.56) \quad D_t^d = E_t \sum_{j=1}^{\infty} NS_{t+j}^d \quad \text{if} \quad \lim_{n \to \infty} E_t D_{t+n}^d = 0.\]

Assuming that the series of debt is represented by a general ARIMA model with constant, Wilcox derives the conditions for sustainability of fiscal policies. His test consists basically in determining whether the discounted debt series is stationary, with a zero mean. Unlike all other tests, this one is based on the transversality condition (2.6) instead of being derived as usually from the IBC.

This type of test was adopted in most empirical studies on the European Union members, such as Corsetti, 1991, Corsetti and Roubini, 1992, Baglioni and Cherubini, 1993, Wickens, 1993, Caporale, 1993b and Uctum and Wickens, 1996, among others. The mixed results obtained can be compared in table A.1 in the Appendix.

The main problem with the methodology introduced by Wilcox is that it requires data on the actual yield on government bonds, which is not easy to obtain, particularly for a long time series. The usual procedure is to choose as proxy the long-term interest rate on government bonds, or to divide the government’s total interest payments by the total debt outstanding, obtaining the ‘interest rate’ used to discount the debt. Both procedures may not capture the true representative interest rate.

Another practical problem, common to all tests based on the series of debt, is that the most adequate measure, the net market value of debt, is not available in the official statistics. To transform the value of the debt from par value, as it is usually presented, to market value is a very difficult task, considering the wide variety of financing sources. For the US there is a series of market debt available, but for the EU studies cited above there is no alternative but to use the par value, which may distort the results of the tests. This question will be resumed in chapter 3.
2.4.2.2. Stationarity of the first-differenced debt

This section presents an alternative to the previous test, avoiding the problems of measurement of the series of debt and interest rate and, at the same time, relaxing the assumption of a constant interest rate. It is not difficult to demonstrate that some of the testing procedures derived in section 2.4.1 cannot be employed without the assumption of a constant interest rate. For example, if the interest rate is allowed to vary, the discount factors will be given by

\[ \delta_{t,n} = \prod_{j=1}^{n} (1 + \rho_{t+j})^{-1} \, . \]

Considering that the net-of-interest surplus observes (2.24), substituting (2.24) in (2.11), with a variable discount factor, gives the following equivalent to equation (2.29).

\[ d_t = E_t \left[ (\delta_{t,1} + \phi \delta_{t,2} + \phi^2 \delta_{t,3} + \ldots) n_{s_{t+1}} + (\delta_{t,2} + \phi \delta_{t,3} + \ldots) \theta(L)e_{t+2} + \right. \]
\[ + (\delta_{t,3} + \phi \delta_{t,4} + \ldots) \theta(L)e_{t+3} + \ldots \]  

The equation above shows that there is no constant cointegrating parameter between the series of debt and seigniorage-augmented primary surplus. The coefficient in \( n_{s_{t+1}} \) is time-varying. This implies that with a variable interest rate, cointegration between \( d_t \) and \( n_{s_{t+1}} \) is no more a condition to ensure and test sustainability of fiscal policies, as proposed before by Trehan and Walsh (1988), MacDonald and Speight (1990) and Haug (1991), for example.

However, as shown in Trehan and Walsh (1991) and Jondeau (1992), stationarity of the total deficit, or of the first-differenced debt, continues to be a sufficient sustainability condition. If the total deficit is stationary it implies, according to the budget identity (2.9), that the stock of debt grows, at most, according to a linear deterministic trend

\[ \rho_t d_{t-1} - n s_t = \Delta d_t = e_t + k \, , \]

where \( k \) is a constant, and therefore

\[ E_t d_{t+n} = E_t \left[ d_t + nk + \sum_{j=1}^{n} e_{t+j} \right] = d_t + nk \, . \]

On the other hand, the series of the variable discount factor \( \delta_{t,n} \) evolves according to
Assuming that $p$ is a stochastic process strictly bounded below by $\gamma > 0$, i.e., that the expected, growth corrected, real rate of interest is positive, $E_t(\rho_{t+n}) > 0$, equation (2.60) shows that the series of the discounted factor decreases in an exponential way. Therefore, from (2.59) and (2.60) it can be observed that $d_{t+1}$ is of smaller order than $\delta_{t,n}$ and therefore it is possible finally to conclude that the transversality condition $\lim_{n \to \infty} E_t \delta_{t,n} d_{t+n} = 0$ is always satisfied.

Applying this test to US fiscal data over the period 1960-84, Trehan and Walsh (1991) found evidence of stationarity of the first-differenced debt series, a sufficient condition to satisfy the IBC. However, this contradicted the results of their cointegration test, on the same data set but assuming a constant expected value of the interest rate, which indicated unsustainability. They therefore conclude that “the government’s budget does not violate intertemporal budget balance, but that the assumption of a constant real rate is at odds with the data” (Trehan and Walsh, 1991: p. 216).

An alternative, less restrictive approach to derive testable implications of the intertemporal budget constraint was proposed by Hakkio and Rush (1991a), and adopted in the most recent empirical studies by, among others, Haug (1995), Quintos (1995), Liu and Tanner (1995) and Crowder (1997), all studying the US. They examine the equilibrium relationship between total expenditures and total revenues of the government, rather than the relationship between the stock of debt and deficits.

The test procedure developed by Hakkio and Rush (1991a) does not require the expected value of the interest rate to be constant. Instead, they assume that the interest rate is stationary, with unconditional mean equal to $\rho$.\(^{39}\) Furthermore, this test also has the advantage, over the Wilcox’ test for example, of using flow variables instead of stock

\(^{39}\) This condition is much weaker than the one assumed in section 2.4.1 of a constant conditional expected value. For example, if the variable follows a stationary ARMA process, its unconditional expected value is constant while its conditional expected value is not. Note that under this assumption, it is not correct to analyse the IBC in nominal terms, since if assuming a stationary real interest rate is still controversial, assuming the stationarity of the nominal interest rate is highly questionable.
variables, which may reduce the risk of measurement errors linked to the correct concept of debt. Rearranging equation (2.8) to give

\[(2.61) \quad d_i = (1 + \rho)d_{i-1} + g_{t-1} - t_i - \Delta m_t - m_{i-1}\left[1 - \frac{1}{(1 + \pi_i)(1 + \psi_i)}\right],\]

where \(g_t = g_t + (\rho_d - \rho)d_{t-1}\), and solving (2.61) forwards, the government's intertemporal budget identity is obtained, equivalent to equation (2.3) but with the variables now expressed as fractions of GDP,

\[(2.62) \quad d_i = \sum_{j=0}^{\infty} \frac{1}{(1 + \rho)^j} \left\{t_{i+j} + \Delta m_{i+j} + m_{i+j-1}\left[1 - \frac{1}{(1 + \pi_{i+j})(1 + \psi_{i+j})}\right] - g^{*}_{i+j}\right\} + \lim_{n \to \infty} \frac{1}{(1 + \rho)^n} d_{i+n}.\]

Assuming that the limit term is zero, it is possible to derive the solvency conditions

\[(2.63) \quad \lim_{n \to \infty} \left[\frac{d_{i+n}}{(1 + \rho)^n}\right] = 0, \text{ and }\]

\[(2.64) \quad d_i = \sum_{j=0}^{\infty} \frac{1}{(1 + \rho)^j} \left\{t_{i+j} + \Delta m_{i+j} + m_{i+j-1}\left[1 - \frac{1}{(1 + \pi_{i+j})(1 + \psi_{i+j})}\right] - g^{*}_{i+j}\right\}.\]

After some mathematical manipulation of equation (2.64), similar to the procedure used to arrive at equation (2.42) in section 2.4.1.3, and therefore not replicated here, the following expression can be obtained

\[(2.65) \quad \Delta t_g - \Delta t_t = \sum_{j=0}^{\infty} \frac{1}{(1 + \rho)^j} \left(\Delta t_{i+j} - \Delta g^{*}_{i+j}\right),\]

where \(t_g\) represents total government expenditures \((g_t + \rho d_{t-1})\) and \(t_t\) stands for total government revenues \([t_t + m_t + m_{t-1}/(1 + \pi_t(1 + \psi_t))]\), as ratios of GDP.

All the variables on the right-hand side of equation (2.65) are expressed in first differences. If all these series in first differences are stationary, i.e., if all the variables in levels are \(I(1)\), the right-hand side of (2.65) is stationary. This implies that the left-hand side, the global deficit, must also be stationary. Since \(t_g\) and \(t_t\) are, by the previous assumption, integrated of order one, they must be cointegrated.
Assuming that both series of total government expenditures and revenues follow a random walk process with drift, Hakkio and Rush (1991a) use the following regression to test for cointegration

\[ t_t = \alpha + \beta t_{tg} + \epsilon_t \]

If \( t_{tg} \) and \( t_t \) are nonstationary, the joint sustainability conditions are that:

A) they are cointegrated, and

B) the parameter \( \beta \) equals unity.

Condition A is a necessary condition to satisfy the intertemporal budget constraint, whereas B is not an indispensable condition, as long as \( \beta \) is positive. To see this, it is necessary to substitute the estimate of \( t_t \) from (2.66) in equation (2.8) and, as before, expand it forward to give

\[ d_{t,n} = \sum_{j=0}^{n} \left( 1 + \rho - \rho \hat{\beta} \right)^{n-j} \left[ (1 - \hat{\beta}) g_{t+j} - \alpha \right] \left( 1 + \rho - \rho \hat{\beta} \right)^{n+1} d_{t-1}. \]

Substituting now on equation (2.63), it becomes

\[ \lim_{n \to \infty} \left\{ \sum_{j=0}^{n} \frac{(1 + \rho - \rho \hat{\beta})^{n-j}}{(1 + \rho)^n} \left[ (1 - \hat{\beta}) g_{t+j} - \alpha \right] \frac{(1 + \rho - \rho \hat{\beta})^{n+1}}{(1 + \rho)^n} d_{t-1} \right\}. \]

As long as \( 0 < \hat{\beta} \leq 1 \), this limit equals zero, i.e., the discounted value of debt equals zero despite the fact that the government will run a permanent total deficit. As can be seen from equation (2.66), the government’s total expenditures will grow at a higher rate than the total revenues.

However, following McCallum (1984), Hakkio and Rush (1991a) have noted that as long as \( \hat{\beta} < 1 \), the undiscounted value of debt will grow without bound, diverging to infinity (as can be proven in equation (2.67) by taking the limit of \( d_{t,n} \)), and the government will

---

40 Condition B is a much stronger condition than A, implying that total expenditures and total revenues must trend together pound-for-pound and therefore the total deficit tends towards a fixed mean.

41 To be more precise, the debt-GDP ratio will only start increasing when the ratio of total expenditures to GDP is above a certain threshold, \( t_{tg} > c_\alpha(1 - \beta) \).
have a larger incentive to default, through repudiation or high inflation. If the investors expect such a default, the government may face some problems in placing its debt on the market. This is why these latter authors have classified $B$ as a 'probably necessary' condition.

2.4.3. The sustainability tests in a stochastic economy

As noted above in section 2.2, the empirical evidence presented, for example, by Abel et al. (1989) indicates that the economies are effectively dynamically efficient. In a deterministic environment, this implies that the interest rate is above the growth rate of the economy, $r > \psi$ or $\rho > 0$. An important consequence from this hypothesis is that the IBC and the transversality condition must necessarily be satisfied in order to ensure solvency.

However, there is also some empirical evidence, notably for the United States, suggesting that the real interest rate has been on average below the growth rate of the economy (for a more recent analysis, in the period 1920-1992, see Ball, Elmendorf and Mankiw, 1998). One consequence is that if in fact $\rho < 0$, the government could actually engage on a Ponzi scheme. This apparent contradiction can only be settled when considering a stochastic economy. As shown above, the efficiency criterion for stochastic economies is different, and it is compatible with a negative $\rho$.

This successive reasoning led Bohn (1991b and 1995) to criticise and reject all tests developed on a nonstochastic economy. In an uncertain world of risky interest rates, none of these tests would be satisfactory. He proposes a new approach to the IBC, with risk-adjusted discount factors, within the stochastic general equilibrium model considered.

42 This latter case represents, basically, the argument of Sargent and Wallace (1981) on the 'unpleasant monetarist arithmetic'.

43 As noted before, the government could continuously, and indefinitely, borrow to pay all interest expenses on accumulated debt. Under these circumstances, the ratio debt-GDP could be continuously decreasing, even with permanent primary deficits. In other words, the country could 'grow itself' out of its debt problem (see Table 2.1, above). The viability and the conditions necessary for the maintenance of these schemes in a stochastic economy are analysed by Bohn (1995) in the Lucas (1978) -type model examined in this section, and by Blanchard and Weil (1992) in an OLG model with incomplete markets.

44 The Abel et al. (1989) efficiency condition states that the cash-flow generated by capital should be greater than gross investment.
below. This approach reconciles the observation of negative growth-corrected real interest rates with the need for the government to obey its IBC.

Suppose there is a market for full contingent claims, where all future government revenues and expenditures can be traded (see Lucas and Stokey, 1983). The price, at period $t$, of a claim at period $t+n$, if the state of nature at that future time is $\alpha_{t+n}$, is represented by $P(\alpha_{t+n})$. These states of nature $\alpha_{t+n}$ have a discrete probability distribution and are taken, at each date, from a set $A_{t+n}$. The government's intertemporal budget identity can now be alternatively represented by

$$D_t = \sum_{n=0}^{\infty} \left[ \sum_{\alpha_{t+n} \in A_{t+n}} P(\alpha_{t+n})NS(\alpha_{t+n}) \right] + \lim_{n \to \infty} \sum_{\alpha_{t+n} \in A_{t+n}} P(\alpha_{t+n})D(\alpha_{t+n}),$$

where, as before, $D_t$ and $NS_t$ are stochastic processes representing, respectively, the government's debt and the net-of-interest surplus. The expression $NS(\alpha_{t+n})$, for example, represents the value of the net-of-interest surplus at time $t+n$ if the state of nature in that period is $\alpha_{t+n}$. Multiplying by $P(\alpha_{t+n})$ yields the market value of the surplus in $t+n$, discounted to period $t$, given the state of nature $\alpha_{t+n}$.

The government's borrowing capacity depends on the investors' willingness to lend. It is therefore necessary to ascertain how the investors will value these claims. For that purpose, it is essential to assume an equilibrium model of the economy. Bohn uses the Lucas' (1978) simple consumption-based model of asset pricing in a pure exchange economy, with identical and infinitely lived individuals, extended to include a government sector. This model basically shows that individuals adjust their consumption and investment plans intertemporally so that the expected discounted value of the returns from an asset will have an equilibrium relation with the ratio of marginal utility from current and future consumption.\(^4\) From the first-order equilibrium conditions it is possible to derive an expression to price assets as a function of consumption.

Suppose that the objective of the representative investor (and simultaneously consumer), in period $t$, is to choose consumption and investment plans so as to maximise the present

\(^4\) Rose (1988) has shown that this model requires similar time-series properties of the growth rate of consumption and the real interest rate. He presents evidence that this condition is not verified in the US.
discounted value of the expected, time-additive utility from consumption over an infinite horizon, according to the intertemporal utility function

\[(2.70) \sum_{n=0}^{\infty} E_t \beta^n U(C_{t+n}),\]

where \(\beta\) is the fixed discount factor \((0<\beta<1)\), reflecting the consumer's subjective rate of pure time preference, \(U(.)\) is an increasing concave utility function, \(C_t\) is the stochastic process representing consumption and \(E_t\) is the expectations operator, conditional on information available in period \(t\).

If this investor wanted to dispose of an extra unit of consumption at period \(t+n\), he would have to give up of \(P(\alpha_{t+n})\) units of consumption today, considering a state in the future of \(\alpha_{t+n}\). Along the optimal path of consumption, the utility lost for not consuming \(P(\alpha_{t+n})\) units today [given by \(P(\alpha_{t+n})U'(C_t)\)] must equal the expected utility gained for the extra consumption in period \(t+n\)

\[(2.71) P(\alpha_{t+n})U'(C_t) = \pi(\alpha_{t+n})\beta^n U'[C(\alpha_{t+n})],\]

where \(\pi(\alpha_{t+n})\) is the probability of occurrence of the state of nature \(\alpha_{t+n}\).

Solving for \(P(\alpha_{t+n})\) gives the marginal rate of intertemporal substitution of consumption between periods \(t\) and \(t+n\), which will henceforth be denoted by \(u_{t,n}\). In the optimum point of consumption, this marginal rate of substitution must equal the rate at which the market is willing to trade consumption intertemporally. Substituting in (2.69), yields the stochastic transversality condition

\[(2.72) \lim_{n \to \infty} E_t \{ u_{t,n} D_{t+n} \} = 0,\]

and the stochastic IBC\textsuperscript{46}

\[(2.73) d_t = \sum_{j=1}^{\infty} E_t \{ u_{t,j} NS_{t+j} \} .\]

\textsuperscript{46} The prices of the contingent claims are proportional to the conditional probabilities and so it is possible to write these expressions in terms of expectations.
These two equations are similar to the usual IBC and transversality condition of a deterministic economy, except that in this model the fiscal variables are not discounted by the interest rate on public debt, but by the marginal rate of intertemporal substitution of consumption $u_{t,n}$ between periods $t$ and $t+n$.

An important consequence of this difference is that, in the stochastic setting, these marginal rates of substitution may be correlated with the debt and surplus variables. Bohn (1995) demonstrates that applying the Euler equation from the consumer's intertemporal first order conditions, linking the marginal rate of substitution with asset returns, $E_t[(1+r_{t+j})u_{t,i}]=1$ (Lucas, 1978), to the above equations yields

\begin{align}
\lim_{n \rightarrow \infty} [\delta_{i,n} \cdot E_t(D_{t+n}) + \text{cov}(u_{t,n}, D_{t+n})] &= 0 \\
\sum_{j=1}^{n} [\delta_{i,j} \cdot E_t(NS_{t+j}) + \text{cov}(u_{t,j}, NS_{t+j})] &= D_i,
\end{align}

where $\delta_{i,n} = \prod_{s=1}^{n} (1 + r_{s+j})^{-1}$. The first term in these equations involves discounting at the risk-free interest rate.

Equation (2.75) shows that the net-of-interest surplus can be negative, on average, as long as it has a sufficient positive covariance with the marginal rate of substitution. The covariance terms are zero only in a deterministic economy or with risk neutral individuals (the marginal utility from consumption is constant). "With risk aversion, the covariances will vanish only if future primary surpluses and public debt are uncorrelated with future marginal utility" (Bohn, 1995: p. 268).

Intuitively, this possibility of a less strict constraint raised by equations (2.74) and (2.75) may be explained as follows. In this stochastic economy, the valuation of future government surpluses, which are expected to cover the present value of the debt, depends on the state of the economy, $\alpha_{t+j}$. For example, assume that the economy is expected to be in a recession at time $t+j$, which causes a relative decline in consumption and therefore an increase in its marginal utility. If the government is expected to present a surplus at time $t+j$ (which will be used to service existing debt), it will be more valued by the investors than if the economy were expected to be in a boom. More generally, a sequence of $NS_t$ in (2.73) will have a higher total value if it is countercyclical than if it is procyclical. This implies that if $NS_t$ is countercyclical, the government can sustain a net-
of-interest deficit on average (or even permanently), discounted by the risk-free interest rate, since the covariance term in (2.75) will be positive.

This argument can be alternatively justified by the following simpler reasoning. A future increase in the net-of-interest deficit will expand private consumption if it raises private wealth in the form of bonds held by the private sector. Given that the marginal utility is a decreasing function of consumption, the result will be a decreasing marginal rate of substitution, $u'$. This reasoning obviously rejects Barro’s (1974) Ricardian Equivalence Theorem, according to which public debt and deficits have no effect on aggregate demand.

The critique on the usual approaches for testing sustainability, and the importance of the positive covariance terms, has been confirmed for the United States by two crucial empirical observations. On the one hand, the net-of-interest budget balance has been negative, on average, for almost the last two centuries, and the debt-GDP ratio has not increased accordingly as expected, which seems not compatible with equation (2.8). However, using equation (2.75) instead, it was demonstrated above that with a sufficiently large covariance, the IBC may be satisfied even if the government continuously run net-of-interest deficits. On the other hand, the growth-corrected real interest rate has been also negative on average, which allows for permanent net-of-interest deficits while debt keeps falling in proportion to GDP.

But there is also empirical evidence which seems inconsistent with the above reasoning. Firstly, Eisner and Pieper (1984) have noted that this relation between the chronic deficits and the stock of debt results from misspecifications of the official accounts. Secondly, the justification of positive covariances succumbs with the observation that both the net-of-interest surplus and consumption are usually procyclical variables: when the economy is growing, consumption and tax revenues increase, while public social expenditures decrease - the surplus increases while the marginal rate of substitution decreases (Hamilton, 1991). This would mean a negative covariance between the two variables, implying that the IBC in (2.75) is even stricter than in (2.11), demanding higher future interest.

47 However, this fiscal policy of maintaining permanent, or even null, primary deficits would be unsustainable. With uncertainty there is a probability in some states of nature of a positive, growth adjusted, interest rate and therefore of an explosive debt-GDP series (see Bohn, 1998). That policy would only be feasible in a dynamically inefficient economy (Diamond, 1965).
net-of-interest surpluses. Therefore, contrary to Bohn's supposition, this analysis does not seem to provide governments with a looser budget constraint.

Nevertheless, the critique of Bohn on all previous tests of the IBC introduces a methodological dilemma. On the one hand, it may not be credible to assume a nonstochastic economy, with a risk-free interest rate on government debt. On the other hand, allowing for stochastic discount rates with risk aversion introduces problems of risk premia, which hinders the execution of an empirical sustainability test.

This problem was first considered by Ahmed and Rogers (1995). Allowing for a stochastic environment, and although imposing an extensive 'but plausible' set of assumptions, they demonstrate that cointegration tests can still be used to test sustainability. The complete demonstration, although relatively simple, is extensive and will be omitted here in the interest of brevity. Taking the first difference from equations (2.72) and (2.73), similarly to the procedure of Hakkio and Rush (1991a), and rearranging, yields

\[
(2.76) \quad G_t + rD_{t-1} - T_t - \frac{\Delta M_t^N}{P_t} = \Delta \sum_{j=1}^{\infty} E_t(\mu_{t,j} NS_{t+j}) - \left[ \lim_{n \to \infty} E_t(\mu_{t,n} D_{t+n}) - \lim_{n \to \infty} E_{t-1}(\mu_{t-1,n} D_{t+n-1}) \right]
\]

where the left-hand side gives the total deficit, equivalent to the change in the stock of debt. From this equation, they demonstrate that to guarantee intertemporal solvency (the limit terms must be zero), a necessary condition is that the first element on the right-hand side is stationary, and therefore the elements on the left-hand side, if nonstationary, must be cointegrated. Imposing a certain structure on the debt process \((D_t = \mu + \lambda t + D_{t-1} + \epsilon_t)\) they subsequently demonstrate that this necessary condition is also sufficient to ensure solvency.

In summary, the results obtained by Ahmed and Rogers (1995) depreciate the objections raised by Bohn to the usual testing procedures. They show that a necessary and sufficient condition to accept the hypothesis of sustainability of the fiscal policy, is the existence of a cointegrating relation between \(G_t, r_tD_{t-1}\) and \(T_t\), with a cointegrating vector \([1 1 -1]\).

Applying this cointegration test to a very long time span of data, from 1792 (1692) until 1992, for the US (UK) fiscal variables, they were unable to reject the null hypothesis of
non-cointegration, concluding in favour of sustainability of fiscal policy in both countries.

In practice, this testable condition is equivalent to the hypothesis that the total deficit is stationary, or that the total revenues and total expenditures are cointegrated. Although starting from a completely different model of the economy, the actual empirical test derived by Ahmed and Rogers is basically equivalent to earlier tests first derived by Trehan and Walsh (1991) and Hakkio and Rush (1991a) for a deterministic economy with a risk-free discount factor.

On the other hand, in a more recent paper, Uctum and Wickens (1996) maintain that this question can be overcome by using fiscal variables discounted to a certain base period, and working directly with these stochastic processes. “This avoids the problem identified by Bohn (1995) of needing to take explicit account of risk when allowing the discount rate to be stochastic” (Uctum and Wickens, 1996: p. 3). Rewriting the intertemporal budget identity with the discount factor defined as \( \delta_{t,n} = \omega_{t} \frac{n}{\omega_{t}} \), so that the factor \( \omega_{t} \), for example, discounts the variable back to a base period 0, yields

\[
(2.77) \quad \omega_{t} d_{t} = E_{t} \left[ \omega_{t+n} d_{t+n} \right] + E_{t} \sum_{i=1}^{\infty} \left[ \omega_{t+i} \left( s_{t+i} + \Delta m_{t+i} + m_{t+i-1} \left( 1 - \frac{1}{(1 + \pi_{t+i})(1 + \psi_{t+i})} \right) \right) \right].
\]

Representing the discounted variables by the superscript \( d \), the transversality condition and the IBC are now identical to the ones previously derived by Wilcox (1989)

\[
(2.78) \quad \lim_{n \to \infty} E_{t} d_{t+n}^{d} = 0, \text{ and}
\]

\[
(2.79) \quad d_{t}^{d} = E_{t} \sum_{j=1}^{\infty} n s_{t+j}^{d}.
\]

They subsequently show that, whether the discounted primary surplus including seigniorage is an endogenous or exogenous process, the conditions for sustainability are that the series of discounted debt and net-of-interest surplus must be stationary with zero mean.

This distinction between endogenous and exogenous surplus was first introduced by Wickens and Uctum (1993) in the context of the current account and the national IBC, and extended by Uctum and Wickens (1996) to the government’s IBC. Recently, Bohn
(1998) resumed this question and proposed a new very simple test which may be used in a stochastic environment and with a negative, on average, growth corrected real interest rate. Intuitively, his test may be explained as follows. Since in some states of nature, permanent net-of-interest deficits may lead to an explosive debt-GDP ratio (see Kremers, 1989), sustainability of the government's fiscal policy implies that a growing debt-GDP ratio must be accompanied sooner or later by an increase in the net-of-interest surplus.

To demonstrate this assertion for example in the context of the Uctum and Wickens (1996) model delineated above, the first step is to apply the discount factor
\[
\omega_i = \prod_{j=1}^{\ell} (1 + \nu_j)^{-1}
\]
to equation (2.8),
\[
d_i = (1 + \nu) d_{i-1} - n s_i,
\]
obtaining the equality
\[
(2.80) \quad d_i^d = d_{i-1}^d - n s_i^d.
\]
Assume that the net-of-interest surplus is endogenous relatively to the accumulated stock of debt, according to the following auxiliary equation
\[
(2.81) \quad n s_i^d = \alpha + \gamma d_i^d + \epsilon_i,
\]
where \(\alpha\) and \(\gamma\) are parameters and \(\epsilon_i\) is a stochastic process representing all other determinants of \(n s_i\). The term \(\epsilon_i\) can be interpreted as the 'autonomous' net-of-interest surplus. Substituting in (2.80) yields
\[
(2.82) \quad d_i^d = \frac{1}{1 + \gamma} d_{i-1}^d - \alpha - \epsilon_i.
\]
The stability of the dynamics in (2.82) requires that \(\gamma\) is positive. In other words, sustainability of the fiscal process is dependent on a significant and positive sensitivity of the net-of-interest surplus to the stock of debt. This response of \(n s_i\) to \(d_i\) may be prompted by a discretionary adjustment by the government, as implicitly assumed in Bohn (1998), or by an automatic adjustment, such as a wealth effect, as suggested by Wickens and Uctum (1993).

The empirical application of this test to the United States showed that with a univariate regression of \(n s_i\) on \(d_i\), the parameter \(\gamma\) is not significant, due to the influence of the economic cycle and of the war years. Therefore, Bohn derives a structural equation for \(n s_i\) from Barro's (1979) tax smoothing model, which implies that \(\epsilon_i\), and consequently \(n s_i\), depends on the level of temporary government spending and on a business cycle indicator.
(the ratio of actual to potential GDP). Incorporating these variables in (2.81), he estimates the resulting equation by OLS, and finds a positive response of the primary deficit to debt, using US data for the period 1916-95 and several sub-periods, an indication of a sustainable fiscal policy.

2.5. Summary

This chapter has surveyed the alternative conceptual approaches suggested in the literature to test the hypothesis of sustainability of fiscal policies. The common starting point of all the approaches is the IBC. This expression is analytically derived from the accounting identity reflecting and synthesising all the financial activities of the government. It shows that solvency requires that the discounted value of the stock of debt must approach zero as time approaches infinity, and therefore the present amount of debt must be exactly matched by the discounted sum of all future government net-of-interest budget balances.

As shown in section 2.2, the IBC only constrains the discounted value of the debt-GDP ratio to go to zero in the limit, it does not impose any limit on the value of the ratio itself. For some authors this constraint is too weak to ensure confidence in the ongoing fiscal process. Although necessary, the IBC is not sufficient to guarantee a more demanding concept of sustainable fiscal policies, as the ones which preserve the stabilisation of the debt ratio.

Section 2.3 presented different measures of this 'strong' concept of sustainability. Basically, these measures differ only by the time horizon in which the debt-GDP ratio is expected to converge to a certain value and stabilise. The theoretical background and the empirical application of these indicators of sustainability have been suggested and explored almost exclusively by international organisations such as the OECD (Chouraqui et al., 1986, Blanchard, 1990 and Blanchard et al., 1990), the IMF (Horne, 1991) and the European Commission (1994), with the aim of developing indicators internationally comparable and relatively easy to compute. These summary measures provide a simple
and quick procedure for fiscal policy evaluation, although the empirical results, as noted in the text, are not always compatible.

The two main weaknesses of this concept of sustainability with a collateral constraint are the reliance on potentially misleading projections of the relevant variables, and the inclusion of implicit normative criteria. It is assumed that the main aim of the fiscal authorities is to stabilise the ratio of debt to GDP, which may be either excessively restrictive or too relaxed.

Due to its arbitrariness, the tests of sustainability with the additional collateral constraint of a bounded debt-GDP ratio will not be directly pursued here. The emphasis of the empirical analysis will be on the more objective notion of sustainability with a transversality condition, surveyed in section 2.4. However, it seems reasonable to acknowledge the potential problems arising from a continuously growing debt ratio. Therefore, the methodology adopted in the empirical tests, and delineated in the next chapter, will include a complementary test on the boundness of the debt ratio. This allows, to a certain extent, a reconciliation of the two major conceptual approaches to test the sustainability of fiscal policies.

Section 2.4 presents a comprehensive review of the tests of sustainability with the transversality condition of no residual discounted amount of debt. These tests basically examine the statistical properties of the relevant time series, and check whether the continuation of the past behaviour of these variables into the future is consistent with the IBC and the transversality conditions.

The various testing approaches differ fundamentally according to the underlying assumptions, notably in what concerns the statistical behaviour of the interest rate. The testing procedures examined in section 2.4.1 begin by imposing the assumption of a constant interest rate. This facilitates the analytical derivation of testable implications of the IBC, but is a very strong assumption that should be avoided in the empirical applications. The tests in the following section relax this assumption, while section 2.4.3 goes further and imposes a complete stochastic model of the economy, with the future variables being discounted by the marginal rate of intertemporal substitution of consumption instead of the risk-free rate on public bonds.
One other factor of differentiation between the testing approaches is the actual variables employed in the econometric estimation. These depend not only on the prior assumptions of the model, but also on the algebraic manipulation of the intertemporal budget equation. Some tests use for example the total surplus, others the net-of-interest surplus, some include the stock of debt, others prefer to examine exclusively flow variables. In theory, it may be equivalent, for example, to test for cointegration between \( G_t, rD_t, T_t \) and \( \Delta M_t^N/P_t \), with cointegrating vector \([1 \ -1 \ -1]\) or between \( TG_t \) and \( TT_t \) with cointegrating vector \([1 \ -1]\). In practice, the conclusions from both tests may differ substantially, depending on the reliability of the data.

A related question concerns the treatment of the seigniorage revenues. Two alternative procedures have usually been followed by the researchers. The first approach disregards completely these revenues, considering them so insignificant as not to matter for the analysis. This assumption may distort the results of the tests, especially in countries where these revenues have been relatively significant at least in part of the sample period considered. The second approach is to include the seigniorage together with taxes in a broad measure of total revenues. This procedure implicitly assumes that the behaviour and effects of both revenue sources are theoretically equivalent in the analysis of sustainability. Both assumptions will be more closely examined in chapter 5.

All these methodological differences suggest the difficulty to summarise, in a concise and meaningful way, the empirical results from all the distinct sustainability tests mentioned above. Table A.1 in the appendix intends to capture systematically some of the main differences in the papers available. In a first glance, the only evident deduction is that the researchers have arrived at completely mixed final conclusions. Besides the differences in methodologies and auxiliary assumptions, there are a variety of other reasons which can justify the extreme difficulty to compare these empirical tests.

The most obvious factor is that they have been applied to a large range of different countries. The US is the only country where a sufficiently large number of tests have been performed allowing more meaningful comparisons. Here, the first conclusion is that the results of the tests are almost perfectly divided, with a slight preponderance in favour
of the sustainability hypothesis. One second general finding, not very surprising, is that the tests performed with data from longer time periods seem more likely to indicate sustainability, while the inverse occurs with shorter and more recent time periods. The papers using samples starting before the first world war (Trehan and Walsh, 1988, Bohn, 1991b and 1998, and Ahmed and Rogers, 1995), some in the eighteenth or nineteenth centuries, all found evidence of sustainability in the US. This obviously results from the long-run horizon characteristics of this type of tests. A longer time span may include the necessary reversal in the fiscal policy. However, this 'rule' has too many exceptions. Studies with as few as twenty-four (e.g., Hamilton and Flavin, 1986, or Trehan and Walsh, 1991) or even fourteen observations (e.g., Wilcox, 1989) are also able to reject the hypothesis of unsustainability.

In what concerns the European Union, there is no unanimity either, although the number of papers concerned is relatively smaller. Table 2.2 summarises the conclusions of the tests where at least one of the six EU members analysed in this study was considered. The only absolute consensus is on the unsustainability of the fiscal policy in Italy. The same, almost undisputed, conclusion was obtained in relation to Belgium.

Many other issues prevent an exact comparison between the sustainability tests. Sample periods selected, data frequency chosen, data sources available, and a large succession of diversified econometric methodologies employed. Among the latter, one important question, often neglected, is the possible existence of structural breaks in the series of debt and surpluses. In a much quoted paper, Perron (1989a) argues that the existence of structural breaks in a series implies that conventional tests are biased towards finding unit roots. This may explain some of the empirical findings of unsustainability of fiscal policies.


50 Although a group of tests on the US, using exactly the same data set collected by Hamilton and Flavin (1986) reached mixed conclusions.
Table 2.2: Summary of the results for the European Union

<table>
<thead>
<tr>
<th>Authors</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grilli, 1989</td>
<td>U</td>
<td>U</td>
<td>?</td>
<td>U</td>
<td>U</td>
<td>S</td>
</tr>
<tr>
<td>MacDonald, Speight, 1990</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>S</td>
</tr>
<tr>
<td>Corsetti, 1991</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>U</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Grilli et al., 1991</td>
<td>U</td>
<td>U</td>
<td>S</td>
<td>U</td>
<td>S</td>
<td>U</td>
</tr>
<tr>
<td>Caporale, 1992</td>
<td>U</td>
<td>S</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>S</td>
</tr>
<tr>
<td>Corsetti, Roubini, 1992</td>
<td>U</td>
<td>S</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>S</td>
</tr>
<tr>
<td>Jondeau, 1992</td>
<td>-</td>
<td>U</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Baglioni, Cherubini, 1993</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>U</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Caporale, 1993b</td>
<td>U</td>
<td>S</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>S</td>
</tr>
<tr>
<td>De Haan, Siermann, 1993</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>S</td>
<td>-</td>
</tr>
<tr>
<td>Heinemann, 1993</td>
<td>U</td>
<td>U</td>
<td>?</td>
<td>?</td>
<td>U</td>
<td>-</td>
</tr>
<tr>
<td>Wickens, 1993</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>U</td>
</tr>
<tr>
<td>Ahmed, Rogers, 1995</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>S</td>
</tr>
<tr>
<td>Caporale, 1995</td>
<td>S</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Roubini, 1995</td>
<td>U</td>
<td>U</td>
<td>?</td>
<td>U</td>
<td>?</td>
<td>S</td>
</tr>
<tr>
<td>Vanhorebeek, Rompuy, 1995</td>
<td>?</td>
<td>S</td>
<td>S</td>
<td>U</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Uctum, Wickens, 1996</td>
<td>U</td>
<td>S</td>
<td>?</td>
<td>U</td>
<td>S</td>
<td>?</td>
</tr>
</tbody>
</table>

A 'S' implies sustainability while a 'U' suggests unsustainability. A '?' indicates inconclusive result while a '-' reveals that the respective test was not performed.

Some of the papers, mostly studying the US particular case, acknowledge this question but either do not examine it empirically, or simply divide the sample in sub-periods and repeat the tests for each. Besides reducing the number of observations available, and therefore the power of the tests, this procedure also depends on the subjective choice of the researcher about the precise date where the break is thought to have happened. This question of structural breaks in the series will be empirically examined in chapter 4. Before the empirical tests, however, the next chapter discusses the necessary issues of methodology and data adopted.
3 - METHODOLOGY AND DATA

This chapter presents and justifies the methodology followed in the empirical part, and describes the alternative data sets employed in the econometric tests. Some additional methodological and data issues, specifically concerned with a particular point in the analysis, will be considered later in the respective sections.

3.1. Methodology adopted

The literature survey in the previous chapter reveals a multitude of alternative approaches available to testing the hypothesis of sustainability of fiscal policies. A comparison of the relative advantages and disadvantages of these testing procedures does not indicate any particular method as the optimal. Ultimately, the preference towards a certain approach depends, among other aspects, on the auxiliary assumptions reckoned to be more credible, the confidence in the time series available, and the sample period and countries considered. The first part of this section reveals and justifies the testing approach followed in the present study, while the second part specifies the main econometric techniques employed.

3.1.1. A sequential testing strategy

The sustainability tests of the next chapter will follow a sequential testing strategy. It covers a comprehensive set of previous approaches and therefore allows more robust conclusions. This sequential strategy, represented schematically in Figure 3.1, will go systematically through the following four main steps:
1) Stationarity tests on total revenues and total expenditures.

Nonstationarity of both series of total revenues and expenditures is a necessary condition to implement the cointegration tests in step three, below. Any other outcome immediately allows a preliminary assessment of sustainability. For example, evidence of stationarity in both series of revenues and expenditures indicates a sustainable fiscal policy, without the necessity of proceeding to the following steps. On the other hand, when the series unequivocally present distinct orders of integration, this may constitute an early indication of unsustainability.

2) Stationarity tests on the total government surplus.

Trehan and Walsh (1988 and 1991) have established that this is a sufficient condition for sustainability, whether a constant or a variable interest rate is assumed. In theory, this condition requires that both series in step one share the same order of integration. Rejection of this hypothesis does not automatically indicate unsustainable fiscal policies, in the wide sense adopted here, as long as the hypothesis in the next step is not rejected.

3) Cointegration test between total revenues and total expenditures.

As demonstrated first by Hakkio and Rush (1991a), this is a necessary and sufficient condition for sustainability when both series are stationary only in first differences (from step one).

4) Hypothesis test of a unitary cointegrating parameter.

Although not strictly necessary for sustainability, this condition prevents the debt-GDP ratio from growing without bound, with potentially disruptive reactions from the financial markets. This hypothetical restriction was implicitly imposed ex ante in step two. Here it will instead be tested for data admissibility.

This particular sequential testing strategy presents several important advantages. In the first place, as mentioned above, it encompasses a large set of different alternative approaches suggested previously in the literature, allowing more convincing test results. The drawback is that it also retains the risk of achieving contradictory results in the alternative tests, and therefore indefinite conclusions.
Variables: total revenues ($t_t$), total expenditures ($t_g$), total surplus ($t_s$), and total debt ($d$)

Secondly, this approach allows the comparison between the two main concepts of sustainability, with a collateral constraint (section 2.3) or a transversality condition (section 2.4). The joint tests in steps three and four are formally equivalent to the test in step two. However, one advantage of adopting a sequential testing strategy of cointegration and hypothesis testing on the parameters is to take into consideration a richer information set. It particularly allows the distinction between the different requirement levels of sustainability.

A third advantage of the proposed testing procedure is that it avoids the unrealistic assumption of a constant interest rate. All these testing conditions are robust to a variable
interest rate specification and even, under certain conditions (Ahmed and Rogers, 1995), to a stochastic economy.

Finally, the variables required to perform the tests are the most adequate given the data available. On the one hand, by using exclusively flow variables, it avoids the dependence on the (inadequate) measure of the outstanding government debt. On the other hand, by focusing on the series of total revenues and total expenditures, it circumvents the need to make necessarily subjective assumptions about the average interest rate on public debt. This would be unavoidable in order to compute the value of the interest expenditures and the primary surplus, not directly available in the official statistics for long periods of time.

3.1.2. The econometric tests

The empirical tests of sustainability rely fundamentally on stationarity and/or cointegration tests on the relevant fiscal variables. The attractiveness of these tests derives mainly from the fact that evidence of a unit root and/or a cointegrating relationship can have direct economic interpretations. Non-rejection of the unit root hypothesis, for example, indicates that exogenous shocks have permanent effects on the variable. The existence of a cointegrating relationship suggests a long-run equilibrium between the variables involved.

To ensure the robustness of the results, several alternative techniques, either of unit roots, cointegration, or causality, will be simultaneously performed and compared. Most of these tests have been extensively explored in the literature, and therefore a very detailed discussion of the procedures will be avoided here. However, references to the underlying theoretical literature are provided throughout this section.

- Unit roots

The univariate tests of unit roots were first proposed by Fuller (1976) and Dickey and Fuller (1979), and became widely known and employed after being applied by Nelson

---

1 The econometric tests were performed mainly with Microfit 4.0 and TSP 4.3A. The software package Rats 4.30 was additionally employed in chapter 4 for seasonal adjustment.
and Plosser (1982) to a large range of macroeconomic data. Several alternatives or extensions to the original Dickey-Fuller (DF) unit root tests have been proposed in the literature, with the objective of increasing its power and deal with a wider class of errors. The most widely used, and which will be adopted here, is the augmented Dickey-Fuller (ADF) test. The standard ADF test is derived from the significance of the parameter $\gamma$ in, for example, the regression

$$\Delta x_t = \alpha + \beta t + \gamma x_{t-1} + \sum_{j=1}^{p} \delta_j \Delta x_{t-j} + \epsilon_t.$$  

One other frequently used test is the Phillips (1987) and Phillips and Perron (1988) nonparametric correction of the DF test, in which allowance is made for possible heteroscedasticity and serial correlation in the residuals. However, this Phillips-Perron (PP) test was shown to exhibit an inferior small sample performance relatively to the ADF test, and therefore should be used only as a complement to other approaches (see for example Schwert, 1989, Campbell and Perron, 1991, Agiakloglou and Newbold, 1992, De Jong et al., 1992, and Liu and Praschnik, 1993).

When testing for unit roots, it is particularly important to take into consideration the sensitivity of the test results to the deterministic components included in the model, namely the trend variable, the lag length employed, the existence of serial correlation in the residuals, and the presence of structural breaks in the series. All these factors will be explicitly considered in the empirical chapters. In the first place, and although most series do not seem graphically to exhibit trended behaviour, a trend variable will always be initially included in the model, and only removed if not significant.

In the second place, particular attention will be devoted to the choice of the lag length $p$. Two alternative methods will be performed and compared, one based on information criteria, the other on a sequential testing approach. In both methods a sufficiently high order must be chosen to initiate the lag selection procedure. The first test includes the number of lags suggested by the ‘Schwarz Bayesian Criterion’ (SBC) and will be denoted by $ADF_{SBC}$. This tends to be more parsimonious than other information criteria, such as the ‘Akaike Information Criterion’ (AIC), and it is shown to be generally consistent if the correct model is included in the set being considered while the AIC is inconsistent (see, for example, Pesaran and Smith, 1998). The second method to choose the lag length was
originally implemented by Perron (1989a), and will be identified as ADF_{ST}. It selects an order $p$ such that the coefficient on the last lagged variable is significant at the 10% level.\footnote{Ng and Perron (1995) show that this method leads to the least size distortions among the several criteria they considered, although at the cost of lower power. In any case, the number of lags is increased if necessary to ensure no serial correlation in the residuals.} Ng and Perron (1995) show that this method leads to the least size distortions among the several criteria they considered, although at the cost of lower power. In any case, the number of lags is increased if necessary to ensure no serial correlation in the residuals.

The power of the standard ADF tests is significantly reduced in the presence of a structural break in the series. Perron (1989a) exposed this problem and proposed a modification to the ADF test which allows for an a priori defined structural break point, chosen by the researcher. However, the assumption of a known break date is highly restrictive and raises questions of data-mining. Two types of alternative approaches have been subsequently developed, intending to endogeneize the date of the break either by using recursive or sequential methods. The basic difference is that the former use changing subsamples of the data while the latter include dummies within the whole sample. For a comparison of some of these methods see Banerjee, Lumsdaine and Stock (1992).

In the empirical chapters, a sequential method proposed by Perron and Vogelsang (1992a) is adopted to assess the relevance of possible structural breaks in the level of a series. This method was recently extended to the case of two structural breaks by Clemente, Montañés and Reyes (1998), which will also be applied. Besides complementing the standard unit root tests, these methods are also very useful to formally identify the dates of major changes in the series, contributing to a better characterisation and understanding of its behaviour.

Even after carefully considering all the above mentioned questions, there remains the problem of the small sample properties of the ADF tests. Cochrane (1991), for example, claims that the unit root tests have low power in small samples. With a small sample, the unit root tests may not encompass a sufficiently long time span to capture the mean reversion necessary to reject the null hypothesis. One obvious solution to minimize this

---

\footnote{The choice of this last lagged term can follow a general-to-specific strategy, beginning with a high order and removing successively the last term until a significant one is found, or a specific-to-general rule, including additional lagged terms until the significance of the last (excluded) term is rejected. Hall (1994) found the former strategy to perform better.}
problem is to maximize the time span of the data set (see section 3.2 below). A more feasible solution, recently proposed, is to use panel data unit root tests to increase the power by expanding the sample size.

Two alternative panel-based unit root testing approaches will be considered. The first was initially proposed by Levin and Lin (1992 and 1993). They examine pooled cross section time series data, testing the null hypothesis that each individual series is nonstationary against the alternative hypothesis that all individual series are stationary. This test imposes the very restrictive assumption of homogeneity of the autoregressive parameter. In other words, when the null hypothesis is rejected, it is assumed that all (stationary) series share the same coefficient of convergence.

Therefore, an alternative test will be simultaneously performed to ensure the robustness of the results. This new test was proposed by Im, Pesaran and Shin (1997) and allows for heterogeneous autoregressive parameters. This IPS test is operationally much simpler than the LL test. Instead of pooling the data, the IPS test computes an average of the individual ADF test statistics of the panel, using demeaned data. Both tests will be presented in more detail, together with their empirical application, in section 4.7.

- Cointegration

After establishing that both series of total revenues and total expenditures are difference-stationary processes, the next step is to test the hypothesis of a cointegrating relationship between them. Granger (1981) introduced the concept of cointegration, acknowledging that a linear combination of nonstationary series could be itself stationary. Several distinct approaches to test this hypothesis have subsequently appeared in the literature.

Gregory (1994) and Haug (1996) compare a variety of cointegration tests, by Monte Carlo methods, and conclude that no test of cointegration completely dominates the others. It is therefore important to report the results of different tests. Two main alternative approaches will be presented in the empirical chapters, based respectively on a single-equation and on a system methods.³

³ Other tests on the existence of an equilibrium long-run relationship between the variables will occasionally be applied in the next chapters, when considered more appropriate given the prevailing
The first approach was initially recommended by Engle and Granger (1987), and consists basically on a direct generalisation of the ADF test to the residuals of the hypothesised linear combination. After estimating, for example, the equation

\[ t_t = \alpha + \beta t_{t-1} + \epsilon_t \]  

by ordinary least squares, an ADF test is performed on its residuals \( t_t - \hat{\alpha} - \hat{\beta} t_{t-1} \).

Cointegration requires that this series is stationary, indicating a long-run equilibrium relationship between the variables. The basic difference relatively to the usual unit root tests outlined above is the set of critical values against which the obtained test statistic is compared.

In alternative to this residual-based test, the empirical chapters will also use the procedure developed in Johansen (1988) and applied in Johansen and Juselius (1990) of maximum likelihood cointegration tests. These are based on the multivariate general representation of the vector error-correction model (VECM)

\[ \Delta y_t = \gamma + \lambda t + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + \psi w_t + u_t \]

where \( \Pi \) is factorised into \( \alpha \beta \), \( \alpha \) is the speed of convergence to the long-run steady state and \( \beta \) is the cointegrating vector. This specification of the multivariate model provides information on the short-run and long-run adjustments to shocks, through the estimates in the matrices \( \Gamma \) and \( \Pi \), respectively. Equation (3.3) is balanced only if the term \( \Pi z_{t-1} \) is stationary. Testing for cointegration basically requires finding the rank of the matrix \( \Pi \). If \( \Pi \) has reduced rank, then there is at least one cointegrating vector present. To assess this, the Johansen approach uses two alternative test statistics, known as the ‘Trace’ and the ‘Maximum-eigenvalue’ (\( \lambda_{\text{MAX}} \)) statistics. The latter tests the null hypothesis of at most \( r \) cointegrating vectors against the alternative of exactly \( r+1 \), while the former tests for at most \( r \) cointegrating vectors against an alternative of at least \( r+1 \) vectors.

---

\( ^4 \) The discussion of which variable should be used as regressand is also postponed to the empirical chapters.
The choice of which and how the deterministic terms are included in the model is a nontrivial issue, and may have considerable influence on the final conclusions of the tests (Johansen and Juselius, 1990). Given the characteristics of the variables used in the empirical chapters, the tests will be performed predominantly in a model with a constant term in the cointegrating vector and no other deterministic elements. The adequacy of this choice may be assessed for example by the likelihood ratio procedure suggested in Johansen (1994)\(^5\)

\[
-2 \ln Q \left( H_i^* \cap H_i \right) = T \sum_{j=1}^{p} \ln \left( \frac{1 - \lambda_j^*}{1 - \lambda_j} \right),
\]

where \(\lambda_j\) are the eigenvalues from the unrestricted estimation and the \(\lambda_j^*\) are the eigenvalues from the restricted estimation, \(T\) indicates the number of observations, \(p\) the number of variables and \(r\) the number of cointegrating vectors. The expression in (3.4) follows the notation in Johansen and compares the model \(H_i^*\) with a restricted constant, with the model \(H_i\) with an unrestricted constant (which implies trended underlying variables), both with no trends. Similar equations can be used to compare other types of models. For example, to test the existence of a trend in the cointegrating vector, comparing between the model with an unrestricted intercept and no trend, \(H_i(r)\), and the model with a unrestricted intercept and a trend restricted to the cointegrating vector, \(H_i^*(r)\), the following statistic is computed

\[
-2 \ln Q \left( H_i \cap H_i^* \right) = T \sum_{j=1}^{r} \ln \left( \frac{1 - B_j^*}{1 - B_j} \right).
\]

The particular statistic obtained in (3.4) is compared with the critical values from the \(\chi^2\) distribution with \(p-r\) degrees of freedom, while the asymptotic distribution of the test in (3.5) is \(\chi^2\) with \(r\) degrees of freedom.

---

\(^5\) Other procedures to choose the type of model include using information criteria, such as the AIC or the SBC, Wald tests on the significance of the deterministic terms, or using the 'Pantula principle', as suggested in Johansen (1992).
Haug (1996) compares several cointegration tests using Monte Carlo simulations, and concludes that the Johansen's maximum eigenvalue test\(^6\) and the Engle and Granger's ADF test reveal the least size distortions. Also using a Monte Carlo study, Gonzalo (1994) shows that the Johansen procedure has better properties than single-equation methods or alternative multivariate methods. It performs better even in the presence of problems of non-normality in the errors or overparameterization, by including too many lags in the model. Kremers et al. (1992), Banerjee (1995) and Zivot (1996) claim that the residual-based tests have low power. Phillips (1991) also argues that the Johansen maximum likelihood tests are generally superior to the residual-based cointegration tests. However, Gonzalo and Lee (1998) recently examined specifically the comparative robustness of the Engle-Granger and Johansen tests and recommend the combined use of both tests. Their recommendation is followed here.

- Causality

As a complement to the central tests of sustainability, chapter 4 will also present tests of Granger-causality between the government's total revenues and expenditures. The advantage of these tests is twofold. On the one hand, they serve as a complementary indication of the existence of a long-run equilibrium relationship between the variables. As shown by Granger (1988), cointegration implies necessarily the existence of a causality effect between the variables, running at least in one direction (the contrary inference may not necessarily hold).

On the other hand, the causality analysis constitutes a natural extension of the sustainability tests. An unsustainable fiscal policy requires an adjustment either on the revenues or on the expenditures side. There is a growing literature on the question of what should be and what actually is the usual succession of policy actions. Does/should the government start by implementing an expenditure plan and then collect the revenues necessary to finance it (spend-tax policies) or does/should it disburse only after assessing the resources available (tax-spend policies)? The Granger-causality tests may help shed some light on this discussion. These tests will also be widely applied in chapters 5 and 6.

\(^6\) Johansen and Juselius (1990) also note that the Trace test has lower power and is probably less reliable.
During the last few years, a considerable amount of research has been focused on the issue of causality in the context of the VAR and/or its restricted form, the VECM. The standard Granger non-causality tests in the context of a VAR in levels are applied, in the particular case of the government budget, to the following representation

\[ u_t = c_1 + \sum_{j=1}^{p} \lambda_{1j} u_{t-j} + \sum_{j=1}^{p} \eta_{1j} t_{g,t-j} + \varepsilon_{1t}, \]

(3.6)

\[ t_{g,t} = c_2 + \sum_{j=1}^{p} \lambda_{2j} u_{t-j} + \sum_{j=1}^{p} \eta_{2j} t_{g,t-j} + \varepsilon_{2t}, \]

where \( u_t \) and \( t_{g,t} \) represent, respectively, the total revenues and total expenditures, and \( c \) is a constant term. With stable VAR models, the tests of 'block non-causality' on one variable are then performed by using a Wald test of the joint significance on the coefficients of all lagged terms of the other variable.

These tests are basically correct provided the 'true' order of integration of the variables involved is zero. However, if the series are integrated and possibly cointegrated, these tests involve nonstandard asymptotic distributions and also nuisance parameters (as shown by Park and Phillips, 1989, Sims, Stock and Watson, 1990, and Toda and Phillips, 1993). This means that the limiting distribution under the null hypothesis has to be simulated for each particular case, which is certainly very laborious and may be even impossible (see, for example, Dolado and Lütkepohl, 1996).

Therefore, in the presence of integrated variables, the standard procedure is to use a VAR specification in first differences, or a VECM in case of cointegration between the variables. In the absence of cointegration, the VAR in first-differences is the appropriate model to estimate (Rambaldi, 1997). However, if the variables are cointegrated and a VAR in first-differences is estimated (to avoid possible 'spurious' results due to the presence of I(1) variables), the model is misspecified, because the error-correction term has been omitted. This results in biased estimates and loss of the long-run information embodied in the original level form of the variables (see Granger, 1988). In this case, the tests must be based on equation (3.3), for example using the error-correction model

\[ \Delta y_t = \gamma + \Pi y_{t-1} + \sum_{j=0}^{p-1} \Gamma_j \Delta y_{t-j} + u_t, \]

(3.7)
where \( y \) is a vector containing \( t_t \) and \( t_g \), and \( \Pi=\alpha\beta \) is, in this particular case, a square matrix with four elements \( \Pi_{ij}=\delta_j^i \) \((i,j=t,tg)\). This allows an additional channel of causality, running from the common trend, or long-run equilibrium relation, which would not be detected by the standard Granger causality test in (3.5). It reintroduces the long-run information lost by differencing.

Suppose we are interested in testing, for example, whether government revenues \( t_t \) are caused by total expenditures \( t_g \). The testing procedure consists basically in assessing the joint significance of all the lagged differenced terms of \( t_g \) and the error correction term (coefficient \( \alpha^g \)), using a conventional Wald test. The resulting statistic is then compared with the critical values taken from the chi-square distribution.\(^7\)

This standard procedure has also been shown to be not entirely correct, since Toda and Phillips (1993) have demonstrated that these tests do not have a chi-square distribution, unless some particular restrictions on the rank of the matrix \( \Pi=\alpha\beta \) (the coefficients of the lagged levels endogenous variables) hold. More specifically in the case analysed here, they show that in the presence of cointegration, the Wald test statistic has only an asymptotically valid chi-square distribution if \( \text{rank}(\beta_g)=1 \) or \( \text{rank}(\alpha_g)=1 \). This results from the fact that the VECM involves the nonlinearity \( \Pi=\alpha\beta \) (see, for example, Yamada and Toda, 1998). Furthermore, the traditional testing procedure is conditioned by a previous sequence of integration and cointegration tests between the variables, and therefore it may suffer from important pre-test biases (given the already mentioned low power of these tests in finite samples and its dependence on nuisance parameters).

These findings motivated the development of several alternative testing procedures. Toda and Phillips (1993), for example, indicate a correct methodology to apply the appropriate tests based on the VECM.\(^8\) However, these alternative procedures are very intricate, and difficult to implement in practice, and have not been widely adopted in the empirical

---


8 A third different approach, known as FM-VAR, was suggested by Phillips (1995), who applies the fully modified least square method originally proposed in Phillips and Hansen (1990) to the VAR model. For a comparison of the performance of the three approaches through Monte Carlo simulations, see Yamada and Toda (1998).
literature (see, for example, Toda and Yamamoto, 1995, Rambaldi and Doran, 1996, and Mills, 1998).

An alternative approach was proposed by Toda and Yamamoto (1995) and by Dolado and Lütkepohl (1996), and it became later known as 'Lag-Augmented Var Approach', abbreviated by LA-VAR. These two papers examine the question of testing restrictions on the parameters of the VAR when there is uncertainty concerning the order of integration of the variables. Their idea was to make Wald tests well behaved by reducing the efficiency of parameter estimators. Since the problems with the traditional tests arise from the singularity of the asymptotic distribution of the least squares estimators of the coefficients of the VAR, the solution was to eliminate the singularity by estimating a VAR model with an order higher than the true model (see, e.g., Dolado and Lütkepohl, 1996).

When there is the suspicion of cointegration between the variables, the procedure is to estimate the VAR model in levels, but increasing it intentionally with one more lag (if the variables involved are I(1), at most) than suggested by the usual order selection criteria. The test is then based on the standard Wald statistic of the significance of all lagged terms except the last one (for applications see, for example, Mills, 1998, and Yamada, 1998). One necessary condition is that the order of integration of the process is not higher than the true lag length of the model. The introduction of the additional lag, which should be zero under the null hypothesis, is necessary to allow the test to become asymptotically chi-square. This enables the use of the standard Wald tests.

One other advantage of this test is that it does not require a pre-test of integration/cointegration of the series, avoiding possible pre-test biases. This may be particularly useful when there are doubts about the existence of a cointegrating relationship between the variables but there are some hypothesis of restrictions on the parameters to be tested. If, for example, the primary intention when applying the VAR model is not to test cointegration but to test the hypothesis of (Granger) causal relations between the

---

9 Choi (1993) presented a similar procedure but for univariate processes.
10 The finding of the order also involves the possibility of some pre-testing bias. However, this is present in the standard procedure as well.
11 The general rule is to introduce as many additional lags as the maximum suspected order of integration of the variables involved.
variables, independently of being or not cointegrated, this approach offers a very simple procedure. However, it is important to keep in mind that the model estimated is only a technical device for testing causality, it is not conceived for inference or forecast.

The main problems with this approach are that it has lower power, relatively to the Toda and Phillips (1993) approach, and it is also less efficient, due to the overfit of the VAR model, i.e., the presence of the additional, redundant terms. However, this depends on the number of variables and lagged terms in the model, as well as on the true order of integration of the series. The shorter the number of variables included and the longer the lag length, the smaller will be the inefficiency caused by the inclusion of the extra lag(s). In this case, the alternative pre-test biases from the unit root and cointegration tests may be more serious. Because of the possible problems of loss of power and inefficiency of this approach, Toda and Yamamoto (1995) suggest the use of this procedure as a complement to the conventional analysis.

3.2. Data Description

This section characterises the data sets employed in the empirical applications of the next chapters, presents the different variables included, and performs a preliminary analysis of their behaviour during the time period considered. The first part justifies the choices concerning the main characteristics of the data, in terms of the countries included, the time span and frequency of the observations, the level of government and the unit of measurement of the variables. The second part presents, in a schematic and concise manner, the actual variables employed, their definitions and sources. Finally, the third part examines informally the evolution of the fiscal variables throughout the sample period, trying to identify the dates and possible causes of the main changes in policy.

---

12 In fact, Yamada and Toda (1998) show that generally the LA-VAR test is less powerful than the other two (VECM and FM-VAR), but it is superior to the others in terms of size. They found a trade-off between size and power of the different tests, which must be considered when interpreting the results of empirical applications, and suggests the pertinence of using more than one approach.

13 Dolado and Lütkepohl (1996) demonstrate this by using Monte Carlo simulations.
3.2.1. Main characteristics of the data sets

The empirical applications in the next chapters will focus on the particular case of the European Union, more specifically on six of its members, Belgium, France, Germany, Italy, Netherlands and the United Kingdom. These countries constitute a more or less homogeneous group, having shared similar objectives of economic integration since the establishment of the European Economic Community (EEC) in 1957. The group includes the founding members of the EEC (excluding Luxembourg for reasons of data availability and relative economic dimension) plus the UK. Although joining the Community only in the early 1970s, and being a member of the Exchange Rate Mechanism (ERM) for less than two years, the UK’s economic policy has been oriented, to a certain extent, towards the same economic objectives of the other countries in the sample. During approximately the last twenty years, the monetary policies in these countries, for example, have been conducted with the common main aim of exchange rate stability vis-à-vis the German mark (Knot, 1998).

This restricted group represents the bulk of economic activity in the EU. Although constituting only two fifths of the current number of members, these six countries represented in 1996 more than four fifths of the total EU’s gross domestic product at market prices.14 On the other hand, in terms of the particular central objectives of this study, this group constitutes a very convenient balanced sample. While one half of the countries (Belgium, Italy and the Netherlands) traditionally present very high debt-GDP ratios by contemporary international standards, suggesting potential sustainability problems, the other half (France, Germany and the United Kingdom) usually displays much lower ratios.

One additional, more technical reason for choosing these particular six countries has to do with data availability. For the set of variables required to perform the econometric tests, these countries provide the widest common time span available for comparable data within the EU. As noted before, the large sample theory of the unit root and cointegration tests recommends the use of a long sample period. The length of the time series may even have a crucial role in the final results of the sustainability tests, given the long-run characteristics of the IBC. A government may present large deficits for considerable

periods of time, as long as this policy is eventually reversed. The wider the time span employed in the tests, the higher the possibility that this policy change is contemplated.

Taking this question into consideration, the tests will be performed on a sample covering the period 1950-1996, the longest time span for which a homogeneous data set is available for all countries involved.\textsuperscript{15} The exact beginning of the sample was chosen following criteria of data reliability and comparability. Although post-war data on public finances can be obtained since 1948 for most countries and variables, there are some reasons to look at the first two years with relative scepticism in certain countries, due to the huge volatility displayed in some series.\textsuperscript{16}

On the other extreme, the sample ends in 1996 essentially because of the suspicions that data for the following years may be deliberately distorted. In fact, 1997 was the relevant year for the calculation of the Maastricht's convergence criteria, and there may have been an exceptional effort on the part of some governments to improve the fiscal indicators, with the aim of joining the other countries in the third phase of EMU.\textsuperscript{17}

The problem is that some of these fiscal adjustments may involve only temporary effects on the measures of the deficit and debt. An improvement of the budget accounts through a wave of privatisations as in France or a cut in public investment as in Germany, for example, has an immediate effect of fiscal retrenchment, without the potential political costs of a tax increase or a reduction of public current expenditures. However, these decisions may yield merely a temporary effect on the deficit, if the expected net returns from those privatised assets or from the curtailed investments exceed, respectively, their sale-value and initial cost.

A third important question when choosing the data set relates to the frequency of the observations. For the variables considered here, there is an inevitable trade-off between the time span of the data and the total number of observations. Quarterly data, for

\textsuperscript{15} One problem of using a long sample period is that it restricts the variables available and a more detailed analysis of their components (as will be apparent in section 3.2.2, below).

\textsuperscript{16} Germany is the exception, since its federal budget for 1950 was the first to be drawn up for the entire federal administration.

\textsuperscript{17} France, for example, obtained significant revenues from privatisations, namely with the sale of France Telecom (0.4 per cent of GDP). The Netherlands changed some accounting rules (\textit{OECD Economic Surveys: Netherlands, 1998, p. 51}). Even in Germany there are claims of accounting manoeuvres to obey the criteria (see, for example, \textit{OECD Economic Surveys: Germany, 1998, pp. 6,59}).
example, increases considerably the total number of data points, but annual data allows a longer time span.

The main data set used in the following empirical chapters favours annual data for two essential reasons. The first concerns the power of the econometric tests. Shiller and Perron (1985) and Perron (1989b) present Monte Carlo evidence that the power of the unit root tests is more responsive to an increase in the time span than on the number of observations. Ng (1995) shows that the power of these tests decreases if the data frequency is increased but simultaneously the time span is reduced.

A similar conclusion emerges when examining the cointegration tests. Hakkio and Rush (1991b) and Lahiri and Mamingi (1995), also using Monte Carlo simulations, conclude that 'time desegregation' does not contribute to an increase in the power of the residual-based cointegration tests. Hu (1996) extends these conclusions to the Johansen cointegration tests. Again, she presents evidence that a larger time span increases the power of the tests much more significantly than a larger number of observations.

A second important reason for preferring annual data are the characteristics of the main variables under analysis. The government's budget is essentially an annual concept (see, for example, IMF, 1986: p.48). For shorter periods of time there may be, for instance, different timings between the payment of the expenditures and the revenues raised, or the financing obtained, to pay for them. Furthermore, higher frequencies include the additional potential problem of stochastic seasonal patterns, which may distort the results of some of the tests. Ghysels and Perron (1993) show that the power of the usual ADF and PP unit root tests may be significantly reduced when using seasonally adjusted data. They conclude that this provides additional support for preferring annual data in the unit root tests. This question of seasonal adjustment will be closely examined in section 4.6.2 where, for a question of completeness, the sustainability tests will be repeated using a quarterly data set.

An additional important choice faced in all empirical studies which include fiscal variables is the appropriate level of government to be considered. Three basic nested levels of government accounts may be considered: the central government, the general government, and the public sector. The general government definition includes the central, regional and local governments, and the social security funds. The public sector
accounts consolidate those with the accounts of all the public enterprises. This last, widest measure of the government accounts can be immediately dismissed for the present study, if only given the unavailability of data prior to the 1970s.\(^\text{18}\) For the two other concepts however, it is not obvious which one is the more relevant measure. The literature has been using both concepts indiscriminately, and sometimes also the public sector definition, although the majority of the researchers choose the central government accounts (see Table A.1 in the appendix).

For reasons of larger data availability, namely a wider range of variables, the main data set used here will consider the central government accounts.\(^\text{19}\) This choice can be further justified with the argument that the central government deficits and debt are more likely to have effects on monetization (Chouraqui, Jones and Montador, 1986) and the interest rate, the object of study of chapters 5 and 6.

Nevertheless, whenever possible the tests will equally be performed with general government data. A comparison between the results from both data sets reveals the relative importance of the regional governments and especially the social security funds in each country. The Maastricht Treaty's fiscal convergence criteria were set in terms of the general government accounts, probably because imposing the constraints on the central government could promote increased deficits and debt in the lower levels of government.

Finally, a choice has to be made regarding the unit of measurement of the variables. The literature on the sustainability of fiscal policies uses nominal and real values of the variables and as ratios of GDP or population (see Table A.1). There are no a priori reasons to believe that this choice affects the results of the tests, as long as all variables are consistently defined. The present study uses variables as ratios of GDP, essentially for the already mentioned reasons of international and intertemporal comparisons.

One question never considered in the papers using the variables in ratios of GDP is the possible boundness of these variables, and its effects on the econometric tests. The ratio

\(^{18}\) Furthermore, the operations and functions of public enterprises are closer to the private sector than to the government. MacDonald and Speight (1990: p. 344), for example, consider that "the lumping together of government and public corporations is undesirable".

\(^{19}\) It is not possible, for example, to obtain reliable data on the general government debt for the entire sample period in all the countries included.
of government primary expenditures and taxes to GDP, for example, can be considered to be bounded below by zero and above by unity. The possibility of bounded variables is not contemplated in the unit root tests. In fact, the true data generating process of these ratios should be modelled as (see Bertola and Drazen, 1993 and Ahmed and Yoo, 1995)

\[ x_t = x_{t-1} + \epsilon_t \quad \text{if} \quad x_\cdot < x_{t-1} < x^*, \]

\[ x_t = x_\cdot + a \quad \text{if} \quad x_{t-1} = x_\cdot, \text{and} \]

\[ x_t = x^* + a \quad \text{if} \quad x_{t-1} = x^*, \]

where \( x_\cdot \) and \( x^* \) represent, respectively, the minimum and maximum values of the ratio \( x_t \) (or 'trigger points', as defined by Bertola and Drazen, 1993), and \( a \) is a positive constant. This, however, introduces nonlinearities in the model which are not taken into account in the tests. 20

However, as claimed by Ahmed and Yoo (1995: p. 990), the standard unit root testing procedure remains valid as long as those limits are not "hit too often in the sample", which seems a reasonable assumption for the fiscal variables. Bohn (1991a: p. 346) also claims that although being bounded, these variables "do not have to be stationary". 21

Furthermore, the variables employed here are not necessarily bounded above by unity, as happened in all the studies cited above. On the one hand, the total expenditures include for example the interest payments on the public debt. On the other hand, total revenues include for example the seigniorage proceeds. And even not considering these monetary revenues, the total amount of resources potentially appropriated by the government equals the sum of GDP and the interest on its debt (Buiter and Kletzer, 1991, and Buiter and Patel, 1992).

---

20 To overcome this question of possibly bounded variables, Allen (1997) suggests the logit transformation (see also Ahmed and Yoo, 1995, and Moosa and Bhati, 1997)

\[ y_t = \log \left( \frac{x_t}{1 - x_t} \right) \]

to obtain an unbounded variable, and repeat the unit root tests on this transformed series. Applying this transformation to the variables here analysed did not change noticeably the previous empirical results.

21 The same is argued by Coakley et al. (1996) in the context of the relation savings-investment rates, and by Hall, Anderson and Granger (1992: p. 120) for the nominal yields to maturity (bounded below by zero). These latter authors argue that since "the statistical characteristics of yields are closer to those of \( I(1) \) series than they are to \( I(0) \) series, […] it is appropriate to treat these yield series as if they were \( I(1) \)".
In sum, three distinct data sets are used in the empirical sections. The main data set includes annual observations of central government data, for the period 1950/96. The second also comprises central government variables, but with quarterly observations during the shorter period 1957Q1/96Q4. Finally, a third data set was compiled, with annual observations of general government data for the period 1950/96 (except for Belgium, 1953/96, and Italy, 1951/96). The next section provides a detailed description of all the variables included in these three data sets.

3.2.2. Variables’ description, sources and adjustments

This section describes the data necessary to construct the variables employed in the empirical tests.\(^{22}\) As noted in the previous section, three data sets were collected, with different characteristics and sources.

The central government’s data, both annual and quarterly, were obtained from the International Monetary Fund’s (IMF) statistical publications. The quarterly series were compiled from several April issues of the International Financial Statistics (IFS), starting from the most recent (April 1998) and moving backwards to obtain the most updated values.\(^{23}\) The same strategy was followed to collect the annual data, taken from the same publication and also from the International Financial Statistics Yearbook (IFSY).

The data on government finance are on a cash basis, and cover the budgetary operations of the central government.\(^{24}\) Operations of other government agencies with individual budgets, and social security funds are excluded. Since the early 1970’s, the annual data on government finance in the IFS issues includes extrabudgetary and social security accounts, not compatible with the definition of central government adopted here. In these cases, if necessary, the annual data is computed from the quarterly values.

---

\(^{22}\) Other occasional data used to illustrate a particular point will be described later, in the respective sections.

\(^{23}\) The April issues have more historical information than the following issues in each year.

\(^{24}\) This is consistent with the analytical presentation of the IBC in chapter two. If instead of being defined on a cash basis, the variables were specified on an accrual basis, an additional term would have to be appended to equation (2.1), representing the accumulation of arrears.
The four tables below present a detailed description of the different series employed to construct the two data sets which use the central government concept. The first two tables describe the data necessary to construct the government finance series in ratios of GDP, which will be used throughout all the empirical chapters. Table 3.1c and Table 3.1d introduce the variables specifically necessary for the tests in chapters 5 and 6, on the effects of the fiscal variables on monetization and the interest rate.

Except where otherwise indicated, the complete series were taken from the IMF publications (the corresponding IMF codes are indicated by a #). In some particular cases, due to the existence of missing values, the series had to be completed or adjusted with data from other sources. The homogeneity of the series was however carefully considered by contrasting different alternative sources. All these adjustments and the origins of the extra values are clearly reported in the respective tables.

It has been suggested that the value of the deficit and debt usually published in the official statistics are not the most illustrative measures of these variables. It may be convenient to make some adjustments to the conventional measures so that they reflect the fiscal situation more accurately. However, as noted above, one drawback of using a long time span of data is that it imposes a limit on the number of additional variables available to perform these adjustments.

The dissatisfaction with the official statistics of the budget balance has centred essentially on two questions, the non-distinction between current and capital expenditures, and the non-adjustment of the deficits to the effects of inflation. Eisner and Pieper (1984, 1988) and Eisner (1984, 1989, 1994) have produced an extensive literature on the correct measurement of the budget deficit in the particular case of the US (see also Buchanan, 1996a, 1996b, and Buiter, 1983, 1985).

The first argument is that the government should follow the accounting practices of private firms, and clearly distinguish between current and capital expenditures. The latter provide potential future revenues for the government, directly through profits or charges, and/or indirectly through the additional taxes from a growing economy. In this sense,

---

25 The exact characterisation of the IMF statistics can be found in all issues of the IFS, and also in the 'Manual on Government Statistics' (IMF, 1986).
Blejer and Cheasty (1991) claim that the deficit on current account is the 'true deficit', considering the capital expenditures to be self-financing.

Table 3.1a: Definition, adjustments and sources of the variables - government finance

Total deficit (-) or surplus (#80): difference between total revenues and total expenditures (#81 + #81z - #82 - #83). This variable can equally be obtained, with the opposite sign, as the sum of the net borrowing by the government (#84) plus the net decrease in government cash, deposits and securities held for liquidity purposes (#87). The IFS do not display primary surpluses, only including interest payments. For the UK it is necessary to subtract the 'Overall adjustment' (#80a), due to a difference in the IFS data sources.

For Belgium, annual data for the period 1950/53 are from the Statistical Yearbook, UN, and quarterly data for the period 1991Q1/93Q2 were supplied by the National Bank of Belgium. For the Netherlands, quarterly data during the period 1957/58 were provided by De Nederlandsche Bank. From January 1992 onwards, all German government finance data include operations in the territory of the unified Germany.

Total revenues: comprises all tax and non-tax government revenues (#81 + #81z) including seigniorage, recorded in non-tax revenues, but excluding proceeds from borrowing.

The variable ‘Revenue’ (#81), as defined in the IFS publications, includes only 'nonrepayable' receipts, i.e., those transactions which do not change the financial asset or liability position of the government. Increases in the value of the government assets are not included because, for example, unrealised capital gains do not reduce its borrowing requirements. 'Grants received' (#81z) are included because they also provide a means of financing expenditure without need to incur in debt.

For Germany, the value in 1950 is from the Statistical Yearbook, UN.

Total expenditures: comprises all government spending (#82 + #83), including interest payments, excluding amortization payments on the stock of debt outstanding.

The variable ‘Expenditure’ (#82) includes the nonrepayable payments made by the government, whether for current or capital purposes. When applicable, ‘Lending minus repayments’ (#83) is also included, comprising government acquisition of claims on others, net of repayments.

For Germany, the value in 1950 was supplied by the Bundesministerium der Finanzen. For Italy, the annual and quarterly values of the government budget balance, from the third quarter of 1991 onwards were computed as the sum, with the opposite sign, of net borrowing in lire (#84b) and in foreign currency (#85b) and monetary operations (#86c). This allows the calculation of the total expenditures (#82). For the third quarter of 1992, the value was provided by the Italian Ministero del Tesoro.

Government debt (domestic, #88, and foreign, #89): stock of direct and assumed debt of the central government, excluding debts of public enterprises, currency issues, debt of the monetary authorities and loans guaranteed by the government (not a source of funds for the government - if it is called to pay them, the value appears in expenditures). It is gross debt, not reduced by the amounts of government claims against others. It is valued at par value, i.e., at the amount the government is obligated to pay when the debt matures.

Data on debt for the UK, after 1960, is from the Annual Abstract of Statistics, CSO. For the Netherlands, the 1950 value is from the Statistical Yearbook, UN.

\[1\] Before 1966, the corresponding IFS codes for the government finance variables were as follows: 81=80a, 82=80b, 83=80c, 86d=82, 88=84, 88b=85a, 89b=85b.

Secondly, there is the argument that the conventionally measured deficits do not take explicitly into account the effect of inflation in reducing the real value of the nominally
denominated public financial liabilities. This amortization of public debt through inflation should be clearly identified and considered in the measurement of public deficits. Since the normal amortization payments are usually not part of the deficit (they are not considered as an expenditure in $G$, see Table 3.1a), the amortization caused by inflation should also not be included in the deficit. Therefore, only the real deficits are relevant to analyse the government’s fiscal position and its effects on the economic activity.

The usefulness of the above suggested adjustments has been questioned by several authors, both in theoretical and empirical grounds (see, for example, Tanzi et al., 1988, De Haan and Sturm, 1995, and Knot, 1996), and a definitive conclusion on the optimal measure has not been reached. Sargent (1987) and Rudin and Smith (1994) claim that the distinction between current and capital expenditures is not necessary for the sustainability analysis, as long as the total surplus includes all sources of revenues. Bohn (1991a) and Vanhorebeek and Van Rompuy (1995) argue that the official budget data is more informative of the government behaviour, assuming that the policymakers are basically influenced by these official, more readily available and publicised, figures (analogous arguments can be found in Artis and Marcellino, 1998).

A similar debate surrounds the measurement of the government’s stock of debt. Again, two central issues have been raised, concerning the choice between gross or net value of debt, and its measurement at par or market value. As noted in Table 3.1a, the available series, employed in the empirical chapters, measures the gross debt valued at par.²⁶

The gross debt includes all interest-bearing financial liabilities of the government, while the net debt basically subtracts its financial assets (e.g. loans to the private sector and shares in enterprises). Uctum and Wickens (1996), for example, consider the net debt to be a more appropriate measure of the fiscal stance. However, one immediate problem with this measure is the difficulty to compute in practice the correct value of all the financial assets included in the highly diversified portfolio of the government. Any attempt remains vulnerable to significant measurement errors. Knot (1996), for example,

---

²⁶ This is also the definition adopted in the Maastricht Treaty, for example (the exact definition of the fiscal variables can be found in article two of the ‘Protocol on the Excessive Deficit Procedure’).
claims that data on gross debt is more reliable than data on net debt (see also Chouraqui, Jones and Montador, 1986).

The second focus of debate is whether to use the readily available par value of debt, or to adjust it in order to obtain the market value. Some authors (Garcia, 1997, Uctum and Wickens, 1996, Elmendorf and Mankiw, 1998, among others) claim that the latter measure is preferable, to reflect the change in the value of the stock of bonds resulting from interest rate movements. Other authors, however, support the use of the par value of debt, namely in the context of the sustainability analysis. Buiter and Patel (1992: p. 183), for example, claim that “par values must be used, since the discount on the debt ensures that [the transversality condition] is always satisfied when market values are used”. Ritchie and Lawton (1993) argue that, in some cases, the market value may send perverse signals about the sustainability of fiscal policies.

The measure of debt considered more appropriate depends also on the perspective from which it is examined. The government is essentially concerned with the interest charges and the amount to be repaid at maturity. The investors are interested in the market values by which it may be sold. As noted in IMF (1986: p. 217), “though a rise in the market value of a fixed-interest security above face value may make its holder feel wealthier, it should not make the debtor feel any poorer, since repayment can be made at par”.

This discussion may be kept at a purely theoretical ground for the US case, where a long series of market values of debt is available to the researchers (assuming it is correctly constructed).27 However, for the EU countries for example, the market value of debt can be found only for Germany and the United Kingdom, and only for a very limited period of time. There have been some attempts to construct a series of market valued debt for the EU countries, in the context of the sustainability tests, but based on strong assumptions. Uctum and Wickens (1996) and Artis and Marcellino (1998) estimate a measure of the market value of debt by multiplying the par value by a market-to-par

---

27 Trehan and Walsh (1988) use both par and market measures of debt and report identical results of the sustainability tests.
index, the implied market price \( D_t^M = (1+r)^{-t} D_t^P \), where the superscripts \( M \) and \( P \) indicate, respectively, market and par values.\(^{28}\)

The unavailability of a more accurate series of net debt at market value is one of the main reasons why the empirical tests of sustainability in chapter 4 will only be performed with flow variables. However, the available series of gross debt valued at par will be used later, for example to evaluate its effects on the interest rate. This would cause some problems of interpretation if both measures follow a significantly different behaviour. However, Artis and Marcellino (1998) conclude that in the EU countries the two measures present a similar evolution. Dwyer (1982) also reports evidence of a high correlation between market and par value of debt. This suggests that, as argued in Bohn (1991a: p. 344), “since the focus is on the long-run, temporary differences between market and par values should not be critical”.

All these fiscal variables are divided by nominal GDP before being used in the empirical applications. Table 3.1b presents the GDP series and the inevitable adjustments necessary to obtain a complete and consistent series for the whole period, especially for quarterly data. In fact, the complete quarterly series of GDP from the IFS is only available for the UK. For the countries and periods where the quarterly data was not available, the annual series was interpolated with a very similar GDP quarterly series from the OECD or, when this was not available, with the industrial production index from the IFS.\(^{29}\) The quarterly variables in this table, unlike all the others, are presented in the sources already seasonally adjusted at annual rates.

The more recent issues of the IMF statistical publications report values of the national accounts for the former West Germany until 1990, and for the unified Germany from 1991 onwards. However, as noted in Table 3.1a, the values of the government finance accounts including the unified Germany start only in 1992. Therefore, the value of the series of national accounts for Germany in 1991 were substituted by values regarding only the former West Germany, in order to make the ratios coherent.

\(^{28}\) Green (1991) suggests a more correct methodology to compute this market-to-par index, but the information required for its construction is not available for the entire sample period and group of countries considered here.

\(^{29}\) For the UK, for example, the quarterly series from the OECD is exactly the same as the one from the IMF.
Table 3.1b: Definition, adjustments and sources of the variables - national accounts

**Gross Domestic Product, current prices (#99b.c):** sum of final expenditures, according to the SNA of the UN.

During the period 1950/52, data for Belgium are from the *Statistics of National Product and Expenditure, OEEC*. For most countries there are large gaps in the quarterly series of this variable. When necessary, the annual values were interpolated with OECD data (Germany: 1960Q1-1978Q4; Italy: 1960Q1-1973Q4) or with the industrial production index of the IFS (Belgium: 1957Q1-1996Q4; France: 1957Q1-1964Q4; Italy: 1957Q1-1959Q4; Netherlands: 1957Q1-1976Q4). The sum of the quarters divided by four equals the annual values.

**Gross Domestic Product, constant prices (#99b.r and #99b.ir):** There are several values missing in most countries. The series were completed using the almost identical OECD data for this variable. In Italy, the change of base between 1959 and 1960 caused an unjustified disruption in the series. To obtain a more consistent sequence, the uninterrupted series from the *IFS Supplement on Price Statistics, IMF*, was used.

**Industrial production index (#66):** this indicator includes the activities of mining and quarrying, manufacturing and electricity, gas and water. It is used to transform the annual values of GDP in quarterly values, when not directly observable.

**Producer or Wholesale Price Index (#63):** registers the evolution in the prices of items at the first important commercial transaction. Used as a deflator of the industrial production index.

---

86

---

1 Before 1966, the corresponding IFS code was 97b. ² There are IFS data available for the period 1960Q1-1967Q4, but it is not seasonally adjusted and therefore it seems preferable to employ OECD data, which uses the same seasonal adjustment program.

Table 3.1c presents the series of long-term nominal interest rates for the different countries, and the measure of the inflation rate used to compute the real interest rate. The relation between the real interest rate and the real growth rate of the economy is crucial in the dynamics of the IBC. However, the approach presented above in section 3.1.1 and applied in chapter 4 to test the sustainability of fiscal policies avoids the problem of choosing the correct proxy for the representative interest rate on government debt. This discussion will therefore be postponed to chapter 6.

The time series of long-run interest rates on public bonds in Germany is not available until 1955, neither from the IFS nor from several other official German sources contacted. Therefore, to allow an homogeneous comparative analysis in chapter 6, the first observations for this series were computed by estimating recursively backwards an OLS regression on the Bundesbank's discount rate (#60), available for the whole period and displaying a very similar behaviour, and future values of the long-run rate.
Table 3.1c: Definition, adjustments and sources of the variables - interest rates

**Long-term government bond yield (#61):** nominal yields to maturity of long-term government bonds. The particular characteristics of this variable in each country, during the period comprised in the sample, are described below.

**Belgium:** [1950/63] - average current yield of 4.0 to 5.5 per cent bonds with five to twenty years maturity; [1964/96] - weighted average yield to maturity of all 5 to 8 per cent bonds issued after December 1962 with more than five years to maturity.


**Germany:** weighted average of all bonds issued by the federal, regional and municipal governments, the railways, the postal system and other institutions, with an average remaining life to maturity of more than three years (four years for bonds included before January 1977).

**Italy:** [1950/77] - yield to maturity of the 5 per cent Government Reconstruction Loan maturing in 1978; [1978/91] - average yields to maturity on bonds with original maturities of fifteen to twenty years; [1991/96] - fixed coupon Treasury Bonds with residual maturities between nine and ten years, based on prices in the official wholesale market.

**Netherlands:** [1950/74] - yield to redemption date (1984) of the 3 ½ per cent bond maturing in June 1998; [1974/77] - unweighted average yield of the two 5 and 5 ¼ per cent government bond issues maturing in 1982; [1977/93] - weighted average yield to maturity of the most recent government bond issues with an average original life of at least ten and a half years; [1993/96] - yield on the most recent ten-year government bond.


**United States:** [1950/77] - unweighted averages of yields, to first call when prices are above par, or maturity, if prices are at par or below, of all fully taxed bonds maturing or callable in twelve years or more (fifteen years or more before 1952); [1977/87] - twenty-year constant maturities; [1987/96] - ten-year constant maturities.

**Inflation rate:** measured as the rate of change of the ‘Consumer Price Index’ (#64), which registers the evolution in the prices of a fixed basket of goods and services acquired by the average consumer. The inflation rate is employed to obtain the *ex post* real interest rate.

Table 3.1d presents some other variables which will be employed in the tests of chapters 5 and 6. The exchange rate of the US dollar in terms of each currency is used to obtain the aggregated variables, the monetary base to compute the seigniorage revenues, and the price of crude petroleum is used as a dummy variable in the causality tests of chapter 6. All these series have been regularly published in the IFS for the entire period required.
provide the necessary time series variables for the entire period: Statistics of National Product and Expenditure, National Accounts Statistics, National Accounts of OECD Countries, and National Accounts vols. I and II.

Table 3.1d: Definition, adjustments and sources of the variables - other variables

- **Exchange Rate** (#rf/wf): measured as US dollars per national currency unit. It is employed on the computation of the relative oil price and the European-wide aggregate measures.

- **Monetary Base** (#14): includes the currency with the private sector and the banks, plus the private and the banker’s deposits in the central bank.

- **Price of Saudi Arabian crude petroleum** (#76aad): weighted average of official state sales prices, in US dollars per barrel. After multiplying by the exchange rate (national currency per US dollar, #rf/wf) and dividing by the domestic consumer price index (#64), yields the price of the barrel in domestic currency and real terms.

The government’s total revenues in this third data set exclude the capital revenues, which are included in the total expenditures as ‘net capital outlays’. For Belgium and Italy there is no government financial data available before, respectively, 1953 and 1951. It is also not possible to collect a homogeneous series of data on general government debt for the countries and time period considered. As before, government data from 1992 onwards is for the whole Germany.

3.2.3. The evolution of the fiscal stance in an historical context

This section examines the main trends in the evolution of the central government variables during the sample period considered in this study. A graphical analysis allows a preliminary identification of the most evident changes in the fiscal process. Some alternative theories advanced to explain these major changes are then analysed. All variables are expressed in ratios of GDP, as in the empirical tests, which facilitates the comparisons between countries and periods of time.

For all the countries considered in the sample, Figure A.1 in the appendix shows a more or less steady increase in the ratios of revenues and expenditures to GDP, indicating a growing share of government activity in the economy during this time interval. It is however possible to identify two major distinct periods in the evolution of these ratios.
During the first half of the sample, both variables followed a similar trend. However, during the 1970s and 1980s the growth in expenditures in all countries clearly exceeds the trend in revenues, resulting in a rapid increase of the deficit. The sudden increase in expenditures coincides more or less accurately with the two oil shocks and other major events described below. This situation seems to have been at least partially reversed during the last decade or so.

The differences in the behaviour of both variables are more clear in Figure A.2, where the series of total surplus are displayed. With the exception of France, a common pattern in all other countries indicates a deterioration of the government finances beginning in the mid 1970s, a slight improvement in the mid 1980s, and a further deterioration in the 1990s. As noted by Von Hagen (1992: p. 19) for example, the evolution of these fiscal variables “can best be characterized by country-specific reactions to common shocks, rather than responses to country-specific shocks”.

The complete transformation in the evolution of public finances in the early 1970s has been attributed to several causes:

i) the most usually quoted cause is the first oil shock (see, for example, Gagales, 1991, Gubian, 1991, and Tanzi and Fanniza, 1996). The argument is that the authorities tried to offset the contractionary effects of the shock through fiscal stimuli, increasing a deficit also affected by the automatic stabilisers. However, Knot (1996), for example, claims that in Belgium and especially Italy, the oil shock was merely an additional negative shock to an already deteriorating situation;

ii) the increasing influence of the keynesian doctrine in the government’s economic policy, suggesting that deficits are politically acceptable and desirable in certain situations (see Buchanan and Wagner, 1977 and Fratianni and Spinelli, 1999, for example);

iii) the deceleration of economic growth in most industrialised countries (Roubini and Sachs, 1989);

iv) the radical transformations in the international exchange rate system, with the end of the Bretton Woods system (e.g., Heinemann, 1998).
The second oil shock in the end of the 1970s also appears to coincide with further problems in the public finances. These were additionally affected by the restrictive monetary policies of the early 1980s, intended to fight inflation, which caused a recession and a large increase in the real interest rates, two important factors in the dynamics of the budget constraint. In fact, the deterioration of the public finances during the 1970s was partly concealed by very low, sometimes even negative, growth-corrected real interest rates. The non-adjustment of the nominal interest rate to inflation reveals the importance of the financial markets in the development of the fiscal stance.\textsuperscript{30}

The evolution of the growth-adjusted real interest rate $\rho_i$ in the six countries of the sample reveals the existence of two clearly delimited periods, before and after the beginning of the 1980s. This can be more formally examined with a simple regression of $\rho$ on a constant and a dummy variable taking a unit value from 1980 onwards and zero otherwise (Table 3.2).

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>-1.2296</td>
<td>-4.6643</td>
<td>-2.4772</td>
<td>-4.6058</td>
<td>-3.9141</td>
<td>-1.8461</td>
</tr>
<tr>
<td></td>
<td>(0.5088)</td>
<td>(0.4748)</td>
<td>(0.6829)</td>
<td>(0.6806)</td>
<td>(0.4961)</td>
<td>(0.4407)</td>
</tr>
<tr>
<td>dummy</td>
<td>5.0199</td>
<td>7.6098</td>
<td>4.1605</td>
<td>6.5300</td>
<td>7.3004</td>
<td>3.9025</td>
</tr>
<tr>
<td></td>
<td>(0.8460)</td>
<td>(0.7894)</td>
<td>(1.1354)</td>
<td>(1.1317)</td>
<td>(0.8250)</td>
<td>(0.7328)</td>
</tr>
</tbody>
</table>

The values in the table are the coefficients from the regression $\rho = \alpha + \beta \text{dummy} + \epsilon$, where the variable \text{dummy} equals unity from 1980 onwards and zero otherwise (standard errors in parenthesis). The sum of both lines indicates the average rate during the period 1980/96.

These results confirm the profound change in the behaviour of the interest rate after 1980, suggesting a significant change also in the sustainability conditions of all these countries. The influence of this structural change will be further examined in the next chapter.

In the second half of the 1980s there was a partial recovery of the total surpluses, although in general they have remained at levels higher than those prior to the 1970s. In the early 1990s the budget balances were again negatively affected by the business cycle, in spite of the significant efforts to comply with the Maastricht criteria.

\textsuperscript{30} This apparent lack of market efficiency may be explained, for example, by the existence of capital controls and financial repression (see, for instance, Giovannini, 1995).
As expected, the flow problems of large deficits were accompanied by the stock problems of a growing government debt (Figure A.3). Since the mid 1970s, the ratio of debt to GDP has maintained a more or less steady growth pattern in most countries of the sample, particularly accentuated during the 1980s. The UK is an exception, with a relatively stable debt-GDP ratio in the last two decades, but its level has been growing in the 1990s and exceeds the relative values of France and Germany. In spite of the recent restrictive fiscal policies, all countries present higher debt ratios at the beginning of EMU than when the Maastricht's debt criterion was agreed.

Several studies have tried to justify this continuous and relatively rapid accumulation of public debt in peacetime, which cannot be entirely explained by the traditional normative arguments of keynesian stabilisation and neo-classical tax-smoothing. During the last decade several studies, mainly from Italian authors, have proposed an alternative explanation based on the political determinants of high deficits and debts. According to these authors, the electoral process in some countries negatively affects the capacity of the governments to solve the fiscal problems. Roubini and Sachs (1989) point to the difficulty of coalition governments to cooperate, while Grilli, Masciandro and Tabellini (1991) suggest that the major problem is the short-term duration of governments in power. Alesina and Perotti (1995) survey these and other political determinants of budget deficits.

More recently, other authors have moved the focus from the political process to the political institutions (see, for example, Von Hagen, 1992, Horstmann and Schneider, 1994, Von Hagen and Harden, 1995, and Alesina and Perotti, 1996). They argue that the budget institutional process, during its various stages of decision, affects the fiscal variables, resulting in a deficit bias.

An alternative explanation is the growing integration of the international financial markets. Goldstein and Woglom (1992), for example, argue that this globalization process facilitates the financing of the deficits, since the governments have a wider market to place its debt, and the interest rates are less sensitive to each country's fiscal situation.

All these arguments can be examined within the game theory framework, which suggests that cooperation is more difficult the larger the number of players and the shorter the time horizon in which they interact.
Whatever the main reasons, the deterioration of the public finances has caused concerns regarding the sustainability of fiscal policies and the negative effects on the economy. This question will be empirically examined in the next chapters.
4. EMPIRICAL TESTS OF LONG-RUN SUSTAINABILITY

4.1. Introduction

This chapter tests the hypothesis of sustainability of fiscal policies, following the sequential testing strategy designed in section 3.1.1. Two fundamental conditions are investigated. The first requires the existence of a long-run equilibrium relationship between the series of total government revenues and expenditures. The second condition goes further, demanding additionally that the ratio of debt to GDP does not grow indefinitely.

A central concern throughout this chapter involves the robustness of the results. On the one hand, alternative econometric techniques are employed in each step along the sequential procedure. On the other hand, the tests are also performed with two additional data sets. The first changes the focus of the analysis to a different level of government, by considering general government data. The second increases the number of data points available, by using quarterly data.

The next section analyses the univariate statistical properties of the main budget variables, with a special emphasis on the total surplus. The unit root tests on the series of total revenues and expenditures may provide earlier indications of the fiscal situation in some countries, but are intended mainly as a pre-test for the cointegration analysis. The test on the total surplus, however, provides a more direct indication of sustainability and stability of the debt ratio. Section 4.3 analyses the possibility of structural changes in the data, and their effects on the results of the nonstationarity tests.

Section 4.4 provides the empirical evidence on the existence of a long-run equilibrium relationship between revenues and expenditures, a necessary and sufficient condition for sustainability of fiscal policies. In addition, it also tests whether the cointegrating parameter equals unity, in which case the non-divergence of the debt ratio would be
ensured. The following section examines the complementary issue of causality in the relation between these variables.

Section 4.6 is devoted to the analysis of the two additional data sets. The first part, using general government data, investigates the influence of the regional governments and the social security budgets in the sustainability analysis of each country. The second part verifies whether quarterly data can supply additional information, or if the behaviour of the series is essentially dominated by stochastic seasonal variations in variables which are conceptually annual.

In section 4.7, two different panel data tests are employed to investigate the hypothesis of a European-wide sustainability of the fiscal policy, an important issue in the process of European integration. The final section summarises the main findings of the chapter.

4.2. The univariate statistical properties of the series

This section investigates the first two steps in the sequential procedure, which consist basically in determining the univariate stochastic properties of the main budget variables. In the first place, the unit root tests are applied to the series of total revenues and total expenditures. This is a necessary preliminary test to allow the cointegration analysis, and may provide some preliminary insights into the fiscal position of the different countries. A more direct analysis is the stationarity test on the total surplus series, performed in the second part of the section.

Table 4.1 displays the results of the unit root tests on the series of revenues and expenditures. In complement to the two main tests discussed in section 3.1.2, two additional unit root tests were performed, the Leybourne's (1995) $ADF_{\text{MAX}}$ and the Phillips-Perron (PP) tests. These may provide complementary information to determine the statistical properties of the series.

---

1 The $ADF_{\text{BC}}$ and $ADF_{\text{ST}}$ tests, where the lag length is chosen, respectively, according to the 'Schwarz Bayesian Criterion' and a 'Sequential Testing' procedure.
The ADF\textsubscript{MAX} test is a simple test which consists basically in choosing the maximum value of the two statistics resulting from applying the standard ADF test to the same series in chronological and time-reversed order,

\[ (4.1) \quad ADF_{\text{max}} = \text{Max}(ADF_f, ADF_r), \]

where \( f \) indicates a forward and \( r \) a reverse realisation of the series. The rationale for this test comes from the observation that, if the series is not symmetrical, these two ADF statistics differ by a non-negligible stochastic term (Leybourne, 1995). This term is an end-effect measure which depends on the difference between the last and the first observation in the series. The two statistics coincide exactly only if these two observations are equal. If the series is nonstationary, as under the null hypothesis of a unit root, this difference will naturally diverge, and the standard ADF test may produce erroneous results.

The statistic obtained in (4.1) is then compared with a set of critical values, provided in Leybourne (1995: p. 565), tabulated using a Monte Carlo simulation in a model with normally distributed errors.\(^2\) He presents evidence that this test is more powerful than the standard ADF test (up to 15% increase in power), whilst sharing similar size properties.

This test is particularly pertinent in situations where there may be a break very early in the series, which seems possible in some of the series tested in Table 4.1 (see figure A.1 in the appendix). In the particular case of the revenue series in France, for example, there seems to be exceptionally low values in the beginning of the series.\(^3\) In these situations, not when the break occurs late in the series, the standard ADF test may spuriously reject the null of a unit root (Leybourne, Mills and Newbold, 1998).

The PP test is a non-parametric adjustment to the ADF test, which is valid under more general assumptions about the error term (Phillips, 1987, Phillips and Perron, 1988). The PP tests in Table 4.1 use a window size of three, chosen according to the formula

---

\(^2\) He only presents critical values for a limited number of sample sizes, but the variations are minimal and not relevant given the values of the statistics obtained here.

\(^3\) The possibility that these unusual low values were due to an error in data collection, or resulted from a (not officially referenced) change of definition in the data sources, was considered and discarded. The values presented by the French national statistical institute (INSEE), for example, are very similar to the IFS series and also display the unusual behaviour in the beginning of the sample.
suggested in Newey and West (1994). The critical values are the same as those used in the standard ADF tests.

Table 4.1: Unit root tests on the series of total expenditures \((tg_i)\) and revenues \((tt_i)\)

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>total revenues</strong> ((tt_i))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF\text{SBC}</td>
<td>-0.8060</td>
<td>-3.4432**</td>
<td>-1.6754</td>
<td>-0.2937</td>
<td>-1.2982</td>
<td>-1.6854</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>ADF\text{MAX}</td>
<td>-0.2447</td>
<td>-2.2190*</td>
<td>-1.6754</td>
<td>-0.2937</td>
<td>-1.1854</td>
<td>-1.6854</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>PP</td>
<td>-0.8669</td>
<td>-4.0866***</td>
<td>-2.0308</td>
<td>-0.2910</td>
<td>-1.4041</td>
<td>-2.1161</td>
</tr>
<tr>
<td><strong>total expenditures</strong> ((tg_i))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF\text{SBC}</td>
<td>-1.4507</td>
<td>-2.1439</td>
<td>-3.6957***</td>
<td>-0.7494</td>
<td>-1.3228</td>
<td>-1.4736</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(1)</td>
<td>(0)</td>
</tr>
<tr>
<td>ADF\text{MAX}</td>
<td>-0.9411</td>
<td>-2.1439*</td>
<td>-3.6957***</td>
<td>-0.7074</td>
<td>-0.3043</td>
<td>-1.3713</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(1)</td>
<td>(0)</td>
</tr>
<tr>
<td>PP</td>
<td>-0.7847</td>
<td>-2.7386</td>
<td>-2.7217</td>
<td>-0.9396</td>
<td>-1.5013</td>
<td>-1.7451</td>
</tr>
</tbody>
</table>

In the ADF\text{SBC} the lag length is chosen according to the ‘Schwarz Bayesian Criterion’. The ADF\text{MAX} is the test proposed by Leybourne (1995) while PP indicates the Phillips-Perron test. In parenthesis is the number of lagged differenced terms used in the regression. In all ADF tests the maximum lag length considered was four. The asterisks (*), (**), and (***), indicate rejection of the null of nonstationarity at the 10, 5 and 1 per cent significance levels, respectively. Critical values of the DF\text{SBC} and PP tests (MacKinnon, 1991): -2.60(10%), -2.93(5%) and -3.58(1%); of the DF\text{MAX} (Leybourne, 1995): -2.14(10%), -2.48(5%) and -3.17(1%).

The ADF\text{ST} test yields exactly the same statistics of the ADF\text{SBC} test in most cases, and therefore was omitted from the table. The only two differences are the inclusion of three lags in the tests of the revenue series in the Netherlands and the United Kingdom, but with no noticeable discrepancies relatively to the ADF\text{SBC} results.

All the results in the table were obtained with models incorporating a constant but no trend. The trend variable was initially included in all models, but later removed because it is not statistically significant. Nevertheless, the tests incorporating the trend produced the same inferences, except in the case of the UK revenues, where the null of a unit root is rejected with the ADF\text{SBC} and PP tests but not with the ADF\text{MAX} test.

The diagnostic tests on the residuals of the estimated models revealed no problems of serial correlation or heteroscedasticity. There were a few problems of non-normality,
which were dealt with, for example, dummy variables without changing the conclusions of the tests.

The tests indicate that most series are integrated of first order. In the case of the series of revenues in France and expenditures in Germany, most tests reject the I(1) hypothesis in favour of stationarity. These results suggest a preliminary conclusion of unsustainability of fiscal policies. However, given the non-unanimity of the test results and the closeness to the critical values, it seems prudent to consider the possibility that these series are also nonstationary, and proceed with the following steps of the testing procedure. If the series of revenues in France is in fact nonstationary with an early break, and if the test on the expenditures in Germany reveals problems in the residuals, the ADFMAX and the PP tests may be, respectively, more reliable. Pain and Thomas (1997) claim that in case of doubt, it is advisable to treat the variables as I(1). Furthermore, it can always be argued that these few rejections of the unit root null are approximately the expected for the usual probabilities of type I error when all series are in fact nonstationary. In any case, these doubts must be carefully considered when analysing the results of the next steps in the testing strategy.

The second step in the sequential approach presents a more direct test of sustainability. A sufficient, but not strictly necessary condition is the stationarity of the series of total surplus. Besides testing sustainability, this condition also examines whether the ratio of debt-GDP will grow without bound. Table 4.2 presents the results of the unit root tests. The ADFMAX was again computed for a question of robustness of the results, given the crucial importance of the stationarity of the total surplus.

All the tests in the table were performed with no trend in the model since this was found not to be statistically significant. However, the tests with models including a trend delivered basically the same conclusions, strengthening the rejection of nonstationarity in the case of Germany [-5.3430 (0)], and weakening the rejection in the UK [-3.3007 (0)]. The standard diagnostic tests did not detect any problem and this is why the PP tests,

---

4 The tests on the first-differenced series, not shown, unanimously rejected the null of nonstationarity at a lower than 1% level, excluding the hypothesis of higher orders of integration.

5 In fact, the ADFST test on the German tG series suggests an optimal lag length of three, which cannot be used due to serial correlation in the residuals.
which have much lower power than the ADF tests with well-behaved residuals, were not presented in the table.

Table 4.2: Unit root tests on the series of total surplus ($ts_t$)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ADF_{BC}$</td>
<td>-1.3384</td>
<td>-2.4347</td>
<td>-4.7637***</td>
<td>-1.3984</td>
<td>-2.1618</td>
<td>-3.2308**</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>$ADF_{ST}$</td>
<td>-1.3384</td>
<td>-2.4347</td>
<td>-4.7637***</td>
<td>-1.1295</td>
<td>-2.1618</td>
<td>-3.2308**</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(1)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>$ADF_{MAX}$</td>
<td>-1.3384</td>
<td>-1.6899</td>
<td>-4.7637***</td>
<td>-1.2112</td>
<td>-1.5900</td>
<td>-2.3751</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
</tbody>
</table>

In the $ADF_{BC}$ and the $ADF_{ST}$ the lag length is chosen, respectively, according to the 'Schwarz Bayesian Criterion' and the 'Sequential Testing' procedure. The $ADF_{MAX}$ is the test proposed by Leybourne (1995). In parenthesis is the number of lagged differenced terms used in the regression. The asterisks (*), (**) and (***) indicate rejection of the null of nonstationarity at the 10, 5 and 1 per cent significance levels, respectively. Critical values of the $DF_{BC}$ and $ADF_{ST}$ tests (MacKinnon, 1991): -2.60(10%), -2.93(5%) and -3.58(1%); of the $DF_{MAX}$ (Leybourne, 1995): -2.14(10%), -2.48(5%) and -3.17(1%).

The unit root tests strongly reject the null hypothesis of nonstationarity of the total surplus in the case of Germany, and are not conclusive in what concerns the UK. In the latter case, the non-unanimity of the tests and the closeness of the statistics to the critical values do not allow a definite conclusion. Further tests are necessary in this case. For all the other countries, no rejection of the unit root hypothesis was found. Only Germany presents a clearly sustainable fiscal policy with a bounded stock of debt as a proportion of GDP.

The robustness of these conclusions will be evaluated with the joint conditions of cointegration with a unit cointegrating term, in section 4.4. Before that, however, the next section searches for possible structural breaks in the series of total surplus, and examines how they may affect the results obtained so far.

---

6 The result for Germany is not compatible with different orders of integration of the revenues and expenditures series, as suggested in some of the unit root tests in Table 4.1, justifying the decision to treat both series as nonstationary.
4.3. The Effect of Structural Breaks

As first suggested by Perron (1989a), the presence of structural changes in the data can bias the results of the standard unit root tests towards a non-rejection of the null hypothesis of nonstationarity. The findings of a unit root in the series of total surplus, in the previous section, may be spurious if they were caused by a certain deterministic structural change. This question is therefore particularly relevant in the context of the sustainability analysis. It may elucidate the reasons for the results of the tests, and indicate the source of unsustainability.

The possibility that the sustainability tests may be influenced by deterministic structural changes affecting the behaviour of the variables has been considered by several authors. In their seminal paper, Hamilton and Flavin (1986) admit that the fiscal conditions may have changed near the end of their sample, affecting the conclusions of the tests. However, they did not investigate this suspicion empirically, due to insufficiency of observations. Wilcox (1989) tests for parameter instability following the method of Chow (1960). Using the same data set as Hamilton and Flavin, he finds evidence of a shift in the structure of fiscal policy in the US, degenerating into an unsustainable position only in the period 1974-84.

Other authors investigate the possibility of a structural break by dividing the full sample in two or more sub-samples, and estimating the model separately for each sub-sample. This is the most popular approach followed in the sustainability literature, and it was used by Hakkio and Rush (1991a), Roberds (1991), Baglioni and Cherubini (1993) Caporale (1993b) and Tanner and Liu (1994), among others. However, this procedure has two major drawbacks. First, the number of data points in each sub-sample may not be enough to ensure sufficiently powerful tests. Second, the dates of the structural breaks were exogenously chosen, according to the author's intuition or a visual inspection of the data.

The first problem was considered in the seminal paper by Perron (1989a), subsequently corrected in Perron and Vogelsang (1993). He presents a model which allows for a one-

---

7 Basically, this method estimates the model for the first \( T-n \) observations of the sample period and then the estimated coefficients are used to calculate out-of-sample forecast errors for the last \( n \) observations of the sample.
time shift in the trend function. The null hypothesis of a unit root with drift and an exogenous structural break is tested against the alternative hypothesis of stationarity around a segmented deterministic trend and/or a shift in the mean. The test consists basically in introducing appropriately defined dummy variables in the standard ADF tests. Perron (1990), corrected in Perron and Vogelsang (1992b), modifies this test for the hypothesis of a changing mean level.

Ahmed and Rogers (1995) applied Perron’s (1989a) test to their sustainability analysis with a very long sample of US and UK data. They try different break-dates, focusing especially on war periods since the nineteenth century, but do not find any evidence contrary to the results obtained with standard unit root tests.

Although allowing the analysis of breaks in the unit root tests using the whole sample, these tests continue to be dependent on the researcher’s subjective choice of the break-date. This risks the possibility of ‘data-mining’. In 1992, several alternative approaches were presented to test the unit root hypothesis in a model including a break in a date endogenously chosen. Some of these approaches were used in the sustainability literature. Tanner and Liu (1994), for example, apply a pre-test procedure, recommended by Christiano (1992), to detect a break in US data. The test finds a break in the fourth quarter of 1981, but confirms the stationarity of the total deficit and therefore the sustainability of the US’s fiscal policy. Makrydakis, Tzavalis and Balfoussias (1999) use an alternative approach, suggested by Zivot and Andrews (1992), where the break point is selected endogenously. Applying this methodology to Greek data, they detect a policy regime change in 1975, which is the deterministic cause of the unsustainability result.

The next section uses an alternative, but closely related, methodology, not applied before to the analysis of sustainability of fiscal policies. It was proposed by Perron and Vogelsang (1992a) for the case of a changing mean, and extended by Perron (1994 and 1997) to the case of trended variables. Like the others mentioned above, this procedure also allows the break-date to be determined endogenously by the model. Section 4.3.2 applies an extension of this test to the possibility of two structural breaks, proposed by Clemente, Montañés and Reyes (1998).

Two alternative conclusions may arise from the application of these unit root tests:
i) The tests continue to be unable to reject the unit root hypothesis. This is evidence that shocks have permanent effects on the government's total surplus, and that the current fiscal policy is in fact unsustainable;

ii) The null hypothesis of nonstationarity is rejected in favour of the alternative of a stationary series with a break in the deterministic component. This suggests that the source of unsustainability is not stochastic.

For the particular group of countries and sample period considered, several specific events are natural candidates to cause a possible structural change in the fiscal variables: the membership of the EEC (UK in 1973, all the others in 1957); the collapse of the Bretton Woods system in 1971; the two oil shocks in 1972/73 and 1978/79; the generalised reversal of the relationship between real interest rates and growth rates in the late 1970s; the membership of the EMS and the ERM after 1979; more recently, the fiscal convergence criteria of the Maastricht Treaty in the early 1990s. It is not possible to isolate a single major event and a particular date where a structural change in the government's total surplus may have occurred. This is an additional reason to let the break-date be chosen by the model, as will happen in the next two sections.

4.3.1. The hypothesis of one break in the series

This section applies the methodology first suggested by Perron and Vogelsang (1992a) to test the presence of a unit root in a series with a level shift. This test follows closely the approach suggested in Perron (1990), but assumes that the break-date occurs at an unknown date, to be determined by the model.

Two distinct models are considered, the 'additive outlier model' (AO), for a sudden change in the mean, and the 'innovative outlier model' (IO), more adequate for a gradual change.\textsuperscript{8} Intuitively, the idea of gradual changes in regime is more appealing. However, the adequacy of the model depends on the characteristics of the data, namely the frequency of observations. For a question of prudence, both models are estimated.

\textsuperscript{8} This designation comes from the literature on the effects of outliers in the econometric models.
The AO model is applied in two consecutive steps. First, the deterministic part of the time series is removed, by estimating the equation

\[(4.2)\quad ts_i = \alpha + \beta DU_i + \delta \tilde{s}_i,\]

where, following the notation in Perron (1989a), \(DU_i\) is a dummy variable equal to one if \(t > T_B\) and zero otherwise, \(T_B\) is the unknown date of the break, and the residuals are represented by \(\tilde{s}_i\). The second step of the procedure consists of using the residuals obtained from equation (4.2) to test for the presence of a unit root, by estimating

\[(4.3)\quad \Delta ts_i = \delta \tilde{s}_{i-1} + \sum_{j=0}^{p} \theta_j D(TB)_{i-j} + \sum_{j=1}^{p} \phi_j \Delta ts_{i-j} + \epsilon_i,\]

where \(D(TB)_i\) equals one if \(i = T_B + 1\) and zero otherwise.

The IO model is applied in only one step, by estimating the equation

\[(4.4)\quad \Delta ts_i = \alpha + \beta DU_i + \gamma D(TB)_i + \delta \tilde{s}_{i-1} + \sum_{j=1}^{p} \phi_j \Delta ts_{i-j} + \epsilon_i.\]

In both models, the test is based on the \(t\)-statistic for \(\delta = 0\), under the null hypothesis of a unit root. An important issue to be considered is the choice of the correct lag length \(p\). Perron and Vogelsang (1992a) compare four different alternative procedures: employ a fixed, pre-determined order, choose according to a \(t\)-test on the significance of the last included lag, or an \(F\)-test on all lags, and choose the order that minimises the \(t\)-statistic for testing the unit root hypothesis. They conclude that this last selection procedure should be avoided, and recommend the strategy of starting with an upper bound and reduce \(p\) successively until the last lagged regressor is significant. Perron (1994) compares this with another alternative procedure to select the lag length, by using information criteria such as the AIC or SBC, and also concludes in favour of the recursive \(t\)-statistic procedure. This was shown to have better size and power properties, and therefore is the approach adopted in this section.

The procedure is systematically repeated for all possible break-dates \(T_B\). The date and the respective test statistic finally considered can then be selected according to two alternative strategies. The first consists of choosing the minimum value of all \(t\)-statistics on the autoregressive coefficient \(\delta\). This is the strategy also followed in the alternative
tests proposed by Banerjee, Lumsdaine and Stock (1992) and by Zivot and Andrews (1992). The second strategy is to choose the date corresponding to the minimum/maximum t-statistic on the coefficient of the dummy variable representing the change in the mean, $\beta$. This method is also followed by Christiano (1992). Both strategies were performed here, but only the former is presented in the table, since they yield similar conclusions (the second strategy indicates break-dates slightly posterior to the ones in Table 4.3). The critical values, higher in absolute value than those in Perron (1989a) for an exogenously chosen break-date, are from Perron and Vogelsang (1992a).

Table 4.3: Unit root tests allowing a structural break in the series ($t_s$)

<table>
<thead>
<tr>
<th>Additive Outlier Model (AO)</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test statistic $t_s$ (lag length $p$)</td>
<td>-2.6773 (1)</td>
<td>-3.0665 (0)</td>
<td>-6.7379*** (0)</td>
<td>-3.7043 (0)</td>
<td>-3.0290 (0)</td>
<td>-4.1230 (1)</td>
</tr>
<tr>
<td>Dummy coef. $\beta$ (stand. error)</td>
<td>-0.0492 (0.0073)</td>
<td>0.0316 (0.0060)</td>
<td>-0.0115 (0.0025)</td>
<td>-0.0700 (0.0068)</td>
<td>-0.0347 (0.0060)</td>
<td>-0.0259 (0.0059)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Innovative Outlier Model (IO)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Test statistic $t_s$ (lag length $p$)</td>
<td>-2.6881 (1)</td>
<td>-3.0106 (0)</td>
<td>-6.5835*** (0)</td>
<td>-3.6633 (0)</td>
<td>-2.9781 (0)</td>
<td>-3.9074 (1)</td>
</tr>
<tr>
<td>Dummy coef. $\beta$ (stand. error)</td>
<td>-0.0119 (0.0061)</td>
<td>0.0096 (0.0053)</td>
<td>-0.0107 (0.0031)</td>
<td>-0.0268 (0.0081)</td>
<td>-0.0106 (0.0054)</td>
<td>-0.0127 (0.0053)</td>
</tr>
</tbody>
</table>

The AO model assumes an instantaneous change in the level of the series while the IO model assumes a gradual change. The number of lags was chosen so that the coefficient on the last included lag is significant at the 10% level of significance (starting from a maximum of three). The (***)) indicates rejection of the null hypothesis of a unit root at the 1% significance level. Critical values (Perron and Vogelsang, 1992a): -4.33/-4.42 (10%), -4.67/-4.76 (5%), -5.20/-5.51 (1%) for the AO/IO models.

The results of the tests in both models reveal a strong rejection of the unit root hypothesis in Germany and a non-rejection in all other countries. This generally confirms the results of the standard unit root tests on the total surplus in section 4.2. In the case of the UK, the previous standard tests yielded contradictory and borderline results. In Table 4.3 the incapacity of rejection of the hypothesis of nonstationarity, and therefore unsustainability, in the UK is more clear.
In addition to the indications of (un)sustainability of fiscal policies, Table 4.3 provides supplementary information concerning the behaviour of the time series of total surplus in the different countries. In this aspect, the results for France stand in clear contrast to all others. Here, the tests indicate a break in 1956, with a positive value of the coefficient of the dummy variable representing the level shift. In fact, as can be observed in Figure A.2 in the appendix, France presented unusually large deficits during the first half of the 1950s, partially recovering afterwards, following the Rueff-Pinay stabilisation plan in 1958 (see for example Gubian, 1991), which aimed at fiscal equilibrium. This test therefore confirms the adequacy of the Leybourne’s (1995) ADFMAX test to the case of France, given the break in the beginning of the series.

For all the other countries, the tests reveal a break in the late 1960s, early 1970s, most of them in 1972/73. The value of the coefficient of the dummy variable representing the change in the mean, and its standard error, suggest a strong and significant negative structural change in the series of total surplus, especially when the AO model is used. These break-dates coincide roughly with the change in the international monetary system and the first oil shock. The exact date of the event which caused the structural change is difficult to assess with these tests. The economic agents may react before the occurrence of the relevant event, if it is known a priori or anticipated, or after the event, if there are time lags in the reaction of the variables.

Although these series have been found to have no deterministic trend, the tests were also performed considering a break in the slope, and a simultaneous break in both the level and slope of the series, following the methodology in Perron (1994 and 1997). The appropriate models differ slightly from equations (4.2)-(4.4). These tests did not yield significantly different conclusions (results not shown). Again, only for Germany do all these variations of the tests reject the null hypothesis of a unit root at the 1% level of significance. The chosen dates of the breaks are different in several cases, but generally do not depart much from the ones indicated in Table 4.3.

---

9 Banerjee, Lumsdaine and Stock (1992) and Zivot and Andrews (1992) suggest that the tests should not be performed in the extreme points of the sample, more precisely outside the interval [0.15T, 0.85T], where T is the number of observations. However, the tests performed here were estimated for all sample points, for completeness. No break-date is outside that interval, although in the case of France it is precisely in the limit. More recently, Perron (1997) proves that the arbitrary exclusion of end-points is not necessary for a correct test.
4.3.2. Multiple structural breaks

The unit root tests in series with one structural break performed in the previous section have shown that, except for Italy (1968) and France (1956), most countries present a structural break in 1972/73. This mid-sample change is the most influential in the behaviour of the series of total surplus. However, there may be other, perhaps less important, structural breaks in the series affecting the results of the unit root tests. Several events, for example related to the integration process in the EU, may have caused these changes.

Lumsdaine and Pappel (1997) have considered the existence of two endogenous structural changes in a model with trended variables, following the methods in Banerjee, Lumsdaine and Stock (1992) for one break. Similarly, Clemente, Montañés and Reyes (1998) extend the sequential break model of Perron and Vogelsang (1992a) to the case of two structural changes in the mean of the series. This latter methodology will be followed here. The null hypothesis being tested is that of a unit root in a process with two structural breaks, against the alternative hypothesis of a stationary process with two deterministic changes in the mean.

There is naturally the question of why not consider the possibility of three or even more break-points (in the limit all observations constitute a break point). In terms of the present study, this seems excessive, given the restricted number of observations available. Besides, the question of the optimal number of break-points to be included in the models, like many other choices discussed above, should be considered in terms of model selection criteria. This has only recently received some attention and no clear conclusions have yet been reached (see the discussion in Maddala and Kim, 1996, Nunes, Newbold and Kuan, 1996, and Kim, 1997).

As before, Clemente, Montañés and Reyes (1998) also consider the AO model for instantaneous changes, and the IO model for gradual changes. In the former case, the first step consists of estimating

\[ \tau_t = \alpha + \beta_1 D U_{1t} + \beta_2 D U_{2t} + \epsilon_t \]
where $D_{U_{i,t}}$ equals one if $t>T_{Bi}$ ($i=1,2$) and zero otherwise, and $T_{Bi}$ is the date of the $i$-th break. After removing the deterministic components of the model, the unit root test is based on the estimation of

\[(4.6) \quad \Delta t_{s_i} = \delta_{s_{i-1}} + \sum_{j=0}^{p} \theta_{o_j} D(TB)_{t_{s_{i-1}}} + \sum_{j=1}^{p} \theta_{o_j} D(TB)_{t_{s_{i-1}}} + \Phi_{j} \Delta t_{s_{i-1}} + \xi_{i},\]

where $D(TB)_{t_{i}}$ equals unity if $t=T_{Bi}+1$ ($i=1,2$) and zero otherwise. In the case of the IO model, the test is performed in just one step, with the estimation of

\[(4.7) \quad \Delta t_{s_i} = \alpha + \beta_{1} D_{U_{i1}} + \beta_{2} D_{U_{i2}} + \gamma_{1} D(TB)_{t_{1}} + \gamma_{2} D(TB)_{t_{2}} + \delta_{s_{i-1}} + \sum_{j=1}^{p} \Phi_{j} \Delta t_{s_{i-1}} + \xi_{i} .\]

Like in the previous section, the two models were estimated for all possible combinations of break points. For a total of $T$ observations, there are $T(T-1)/2$ possible combinations to be considered for each model. In both models the statistic of interest for the test is the minimum $t$-statistic for the significance of the autoregressive parameter $\delta$. This is compared with the critical values tabulated by Clemente, Montañés and Reyes (1998) for finite samples (their values for a sample of 50 observations are used here). The results are presented in Table 4.4.

Even after allowing for the possibility of two changes in the mean of the time series of total surplus, the null hypothesis of a unit root can again only be rejected in the case of Germany. Although the conclusions of the unit root tests do not advance any additional evidence besides confirming earlier results, Table 4.4 provides however other interesting information.

The coefficients of the two dummy variables are generally highly significant (the exception is the second dummy for the UK). This suggests the existence of, at least, three differentiated periods in the evolution of the fiscal deficits in these countries. There is in general a first period of low deficits until the mid 1970s, followed by a more or less extensive period of high deficits, and finally a recovery after the mid 1980s, although to levels above those registered in the first period.

Two main exceptions are worthwhile reporting. In Italy, the second, middle period is more extensive and is characterised by higher deficits than in any other country. It starts in the late 1960s and terminates only in the early 1990s, presumably influenced by the
pressures to achieve the fiscal consolidation required to participate in the last stage of EMU.

Table 4.4: Unit root tests allowing two mean shifts ($t_{ts}$)

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Additive Outlier Model (AO)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test statistic $t$ &amp; (lag length $p$)</td>
<td>-4.7634 &amp; (1)</td>
<td>-5.0513 &amp; (0)</td>
<td>-6.2344** &amp; (2)</td>
<td>-4.2033 &amp; (0)</td>
<td>-4.0908 &amp; (3)</td>
<td>-4.7822 &amp; (3)</td>
</tr>
<tr>
<td>Dummy coeff. $\beta_1$ &amp; (stand. errors)</td>
<td>-0.0764 &amp; (0.0069)</td>
<td>0.0386 &amp; (0.0054)</td>
<td>-0.0124 &amp; (0.0033)</td>
<td>-0.0754 &amp; (0.0060)</td>
<td>-0.0486 &amp; (0.0064)</td>
<td>-0.0298 &amp; (0.0066)</td>
</tr>
<tr>
<td>$\beta_2$ &amp; (stand. errors)</td>
<td>0.0407 &amp; (0.0082)</td>
<td>-0.0165 &amp; (0.0040)</td>
<td>0.0040 &amp; (0.0036)</td>
<td>0.0249 &amp; (0.0121)</td>
<td>0.0214 &amp; (0.0075)</td>
<td>0.0102 &amp; (0.0081)</td>
</tr>
<tr>
<td><strong>Innovative Outlier Model (IO)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test statistic $t$ &amp; (lag length $p$)</td>
<td>-4.3803 &amp; (1)</td>
<td>-4.9336 &amp; (0)</td>
<td>-6.6497*** &amp; (2)</td>
<td>-3.7591 &amp; (2)</td>
<td>-4.1854 &amp; (3)</td>
<td>-4.7554 &amp; (1)</td>
</tr>
<tr>
<td>Dummy coeff. $\beta_1$ &amp; (stand. errors)</td>
<td>-0.0433 &amp; (0.0109)</td>
<td>0.0215 &amp; (0.0057)</td>
<td>-0.0273 &amp; (0.0049)</td>
<td>-0.0354 &amp; (0.0089)</td>
<td>-0.0199 &amp; (0.0063)</td>
<td>-0.0186 &amp; (0.0062)</td>
</tr>
<tr>
<td>$\beta_2$ &amp; (stand. errors)</td>
<td>0.0298 &amp; (0.0077)</td>
<td>-0.0127 &amp; (0.0033)</td>
<td>0.0128 &amp; (0.0039)</td>
<td>0.0290 &amp; (0.0103)</td>
<td>0.0137 &amp; (0.0059)</td>
<td>0.0036 &amp; (0.0064)</td>
</tr>
</tbody>
</table>

The AO (IO) model assumes an instantaneous (gradual) change in the level of the series. The number of lags was chosen so that the coefficient on the last included lag is significant at the 10% level of significance (starting from a maximum of three). The asterisks (***) and (**) indicate rejection of the null hypothesis of a unit root at, respectively, the 1% and 5% significance levels. Critical values (Clemente, Montañés and Reyes, 1998): -5.37/-5.52 (10%), -5.70/-5.88 (5%), -6.50/-6.55(1%) for the AO/IO models.

France is an even more particular case, with the evolution of the total surplus looking almost as a mirror image of all the others: high deficits in the two extreme periods and low deficits or even surpluses in the middle. This is formally shown by the sign of both dummy variables representing the level shifts. In France they always display completely opposite signs comparing with all other countries: a positive sign of the first dummy, suggesting a change to more balanced budgets, and a negative sign in the second dummy, indicating a deteriorating fiscal situation. A possible explanation for the very high deficits in the beginning of the 1950s is the reconstruction costs from World War II, augmented by the subsequent wars in the French colonies. This first period ends with the Rueff-
Pinay stabilisation plan mentioned above. On the other hand, the deterioration of the public finances in the early 1980s can be explained by the expansionary fiscal policies followed after the election of F. Mitterrand in 1981, and his designation of a socialist government whose main aim was to solve the problem of a rising unemployment rate. These policies, mainly impelled by political and social reasons, contrast markedly with the contractionary policies followed at the time by most EU countries, with different fiscal perspectives.

For all countries, the coefficient of the first dummy variable is always higher in absolute value and more significant than the coefficient of the second dummy, revealing the higher magnitude of the first, usually negative, break. Furthermore, the date of the first break coincides in general with the date detected by the previous test of section 4.3.1, where only one change was considered. This may help explain why, in spite of the recovery of the 1980s and 1990s, the fiscal situation is still considered unsustainable in most countries. The recovery was not sufficient to restore the government finances to levels previous to the 1970s. The negative shock in the early 1970s seems to be still affecting the public finances.

In sum, all the results so far suggest that only Germany presents a sustainable fiscal policy, with an expected stable trajectory of the debt-GDP ratio. For all other countries this strong condition of stability of the debt ratio is not verified. As noted before, this implies potential difficulties for the governments to market their debt in the future, particularly in those countries already enduring very high debt ratios. However, these results do not necessarily indicate unsustainable fiscal policies. Sustainability does not strictly require stationary total surpluses and stable debt-GDP ratios, as long as the total revenues and expenditures are cointegrated. This condition is examined in the next section.
This section is reserved to the analysis of steps 3 and 4 of the sequential testing strategy designed to evaluate the sustainability of fiscal policies. As noted in chapter 3, two alternative tests of cointegration are performed, the Engle and Granger (1987) single-equation test, which was the first formalisation of Granger's (1981) concept of cointegration, and the Johansen (1988) multivariate approach.

The Engle-Granger two-step procedure tests the existence of a unit root in the residuals from the cointegrating regressions

\[ mt_t = \alpha^{x} + \beta^{x}tg_t + \epsilon^{x}_t \quad \text{or} \quad tg_t = \alpha^{y} + \beta^{y}mt_t + \epsilon^{y}_t, \]

where the superscripts identify the series used as dependent variable. Since the parameters are allowed to be determined by the regression, not imposed \textit{a priori} as in the previous sections, the residuals from (4.8) are not a measure of the total surplus/deficit. Instead, the residuals from these regressions can be interpreted as a 'budgetary disequilibrium', representing short-term deviations from a possible long-run equilibrium relationship between total revenues and expenditures.

The choice of which of the two representations in (4.8) should be used in the tests is usually not innocuous in finite samples, both in terms of the estimates and of the unit root tests on the residuals. The results are only robust to the choice of regressand when the \( R^2 \) from the regression is sufficiently close to unity so that \( \beta^x \equiv 1/\beta^y \) (since \( \beta^x \beta^y = R^2 \)). Given that \( R^2 = 1 \) is an unrealistic situation, it is necessary to pay some attention to this question, which has been overlooked in most sustainability tests using the Engle-Granger methodology.

Ng and Perron (1997) examine this normalisation issue using Monte Carlo simulations of single-equation estimation of cointegrating vectors in a two-variable model. They show that "least-squares estimates have poor finite sample properties when normalised in one direction, but are well behaved when normalised in the other" (p. 53). They suggest using as regressand the less integrated variable, which may be indicated by comparing the
spectral density at frequency zero of both first-differenced variables (the smallest value indicates the regressand).\footnote{The value of the spectral density function at zero frequency expresses the long-run properties of the variable, it gives a measure of persistence of shocks. Higher values suggest more persistent effects of shocks, and therefore more integrated variables.}

Table 4.5 presents the results of the Engle-Granger cointegration tests. Both representations of the cointegrating equation in (4.8) were estimated. In the bottom of the table, the values of the spectrum at frequency zero are displayed to give an indication of which regression should be preferably considered.

Table 4.5: Engle-Granger cointegration tests

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta t = \alpha + \beta t_g$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF test</td>
<td>-2.0874</td>
<td>-3.4808**</td>
<td>-3.1808*</td>
<td>-1.6684</td>
<td>-4.2896***</td>
<td>-3.7867**</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>parameters:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0765</td>
<td>0.2208</td>
<td>0.0995</td>
<td>0.0354</td>
<td>0.1002</td>
<td>0.1125</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0197)</td>
<td>(0.0088)</td>
<td>(0.0081)</td>
<td>(0.0109)</td>
<td>(0.0200)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.5663</td>
<td>-0.0506</td>
<td>0.2496</td>
<td>0.6149</td>
<td>0.6193</td>
<td>0.6147</td>
</tr>
<tr>
<td></td>
<td>(0.0251)</td>
<td>(0.0848)</td>
<td>(0.0590)</td>
<td>(0.0274)</td>
<td>(0.0341)</td>
<td>(0.0564)</td>
</tr>
<tr>
<td>$\Delta t_g = \alpha + \beta t_{\Delta t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>parameters:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-0.1001</td>
<td>0.2636</td>
<td>-0.0075</td>
<td>-0.0298</td>
<td>-0.1047</td>
<td>-0.0360</td>
</tr>
<tr>
<td></td>
<td>(0.0178)</td>
<td>(0.0544)</td>
<td>(0.0368)</td>
<td>(0.0144)</td>
<td>(0.0232)</td>
<td>(0.0358)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.6225</td>
<td>-0.1550</td>
<td>1.1398</td>
<td>1.4921</td>
<td>1.4210</td>
<td>1.1797</td>
</tr>
<tr>
<td></td>
<td>(0.0719)</td>
<td>(0.2599)</td>
<td>(0.2694)</td>
<td>(0.0667)</td>
<td>(0.0782)</td>
<td>(0.1083)</td>
</tr>
<tr>
<td>Spectral density at frequency zero</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta t_t$</td>
<td>1.4088</td>
<td>0.4583</td>
<td>0.5271</td>
<td>1.2453</td>
<td>0.9757</td>
<td>0.3177</td>
</tr>
<tr>
<td>$\Delta t_{tg}$</td>
<td>1.9080</td>
<td>0.4718</td>
<td>0.1591</td>
<td>1.0663</td>
<td>1.2187</td>
<td>0.3887</td>
</tr>
</tbody>
</table>

The number of lags used in the ADF test (in parenthesis below the corresponding statistics) is suggested by the SBC. The asterisks (***), (**), and (*) indicate rejection of the null hypothesis of no-cointegration at, respectively, the 1%, 5% and 10% significance levels. Critical values (MacKinnon, 1991): -4.1433 (1%), -3.4716 (5%) and -3.1374 (10%). In parenthesis below the coefficients are the standard errors. The values of the spectral density function were obtained using a Bartlett lag window of size $2^\frac{1}{2}T$.\footnote{The value of the spectral density function at zero frequency expresses the long-run properties of the variable, it gives a measure of persistence of shocks. Higher values suggest more persistent effects of shocks, and therefore more integrated variables.}
In the cases of Belgium and Italy, the Engle-Granger residual-based approach clearly reinforces the conclusions of unsustainability obtained with the tests on the 'restricted' cointegration relation of previous sections. For the Netherlands and the UK, on the contrary, revenues and expenditures seem to cointegrate when the cointegrating parameters are allowed to be determined by the regression (although in the latter case the hypothesis of a unit root in the residuals is only marginally rejected). For the other two countries, Germany and France, the results are dependent on the direction of normalisation, and therefore deserve a special attention.

Following the suggestion of Ng and Perron (1997), the variables should be considered cointegrated in both countries. However, it is imperative to distinguish both situations. On the one hand, while the hypothesis of no cointegration is rejected in both regressions for Germany, at the 1% and 10% levels of significance levels, it is only rejected in one regression for France, and very close to the 5% level. On the other hand, the parameter $\beta$ in the French regressions display a negative sign (and is non-significant), suggesting an inverse relation between revenues and expenditures. And, according to Ng and Perron, only one of the two point estimates has very poor properties.

Given the sensitivity of the results to the normalisation of the variables, other approaches should be used to test the existence of cointegration between the series of revenues and expenditures. The multivariate approach of Johansen (1988) has the advantage of considering all variables as explicitly endogenous, avoiding the choices faced above.

Table 4.6 presents the results of the Johansen testing procedure. The model employed in the tests includes a constant in the cointegrating vector and no other deterministic terms. According to the indication of the Schwarz Bayesian Criterion (SBC), a VAR of order two is employed for the Netherlands, and of order one for all other countries.11

The null hypothesis of no cointegration is rejected only for Germany and the Netherlands, reinforcing the strong rejections obtained with the Engle-Granger procedure. In the cases of Belgium and Italy, Table 4.6 unequivocally confirm all previous findings of unsustainability of the fiscal policies. For France and the UK, the results of the different

11 For Belgium, although all selection criteria choose order one, the Lagrange multiplier test of residual correlation shows that in one of the equations of the VAR it is possible to reject the hypothesis of no serial correlation at the 5% level. This problem disappears if the order is increased, without altering qualitatively the conclusions of the tests.
tests are contradictory. However, given the very marginal results of the previous tests, and the superior properties of the Johansen tests (see section 3.1.2), it seems reasonable to include these two countries in the group of EU members currently following unsustainable fiscal policies.

Table 4.6: Johansen cointegration test

<table>
<thead>
<tr>
<th></th>
<th>H₀</th>
<th>H₁</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ_max test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r = 0</td>
<td>r = 1</td>
<td>6.2977</td>
<td>11.1973</td>
<td>20.0808**</td>
<td>10.9491</td>
<td>17.9293**</td>
<td>13.5289</td>
<td></td>
</tr>
<tr>
<td>r ≤ 1</td>
<td>r = 2</td>
<td>4.3051</td>
<td>4.9036</td>
<td>3.1953</td>
<td>1.7886</td>
<td>2.9744</td>
<td>2.3182</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trace</td>
<td></td>
</tr>
<tr>
<td>r = 0</td>
<td>r ≥ 1</td>
<td>10.6027</td>
<td>16.1008</td>
<td>23.2761**</td>
<td>12.7377</td>
<td>20.9037**</td>
<td>15.8470</td>
<td></td>
</tr>
<tr>
<td>r ≤ 1</td>
<td>r = 2</td>
<td>4.3051</td>
<td>4.9036</td>
<td>3.1953</td>
<td>1.7886</td>
<td>2.9744</td>
<td>2.3182</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In parenthesis is the order of the VAR. The asterisks (**) indicate rejection of the null hypothesis of no cointegration at the 5% level of significance. Critical values (Pesaran, Shin and Smith, 1997): 15.87/20.18 (5%) and 13.81/17.88 (10%) for the λ_max/Trace tests and the null hypothesis of r=0, and 9.16 (5%) and 7.53 (10%) for both tests of r≤7.

Sustainability, as argued above, does not guarantee a bounded debt-GDP ratio. The final step in the sequential testing approach adopted here is to test this boundness condition with a test on the parameters of the cointegrating relations found in Table 4.6. For the countries where it is not possible to reject the hypothesis of a cointegrating relationship, Table 4.7 presents the values of the cointegrating parameters (and the corresponding standard errors), together with the results of a likelihood ratio test on the restriction β=1.

These results confirm the conclusions from the unit root tests on the series of total surplus, when the restriction β=1 is imposed a priori. In spite of presenting a sustainable fiscal policy, the ratio of debt in the Netherlands may rise indefinitely, and the government may experience difficulties to roll over its debt, faced with the market discipline from the financial markets. Only Germany satisfies both major notions of sustainability, obeying not only the transversality condition but also the collateral constraint of a bounded debt-GDP ratio.
Table 4.7: Tests on the parameters of the error correction terms

<table>
<thead>
<tr>
<th></th>
<th>Germany</th>
<th>Netherlands</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-t_{lt} + \alpha + \beta t_{gt}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameters:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.0240</td>
<td>0.0910</td>
</tr>
<tr>
<td></td>
<td>(0.0366)</td>
<td>(0.0159)</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>0.7602</td>
<td>0.6461</td>
</tr>
<tr>
<td></td>
<td>(0.2469)</td>
<td>(0.0490)</td>
</tr>
<tr>
<td>LR test of restrictions: $\hat{\beta} = 1$</td>
<td>0.5524</td>
<td>10.8270***</td>
</tr>
<tr>
<td>ECT coefficient in equation: $D_{tt}$</td>
<td>0.0804</td>
<td>0.5303</td>
</tr>
<tr>
<td></td>
<td>(0.0694)</td>
<td>(0.1542)</td>
</tr>
<tr>
<td>$D_{tg}$</td>
<td>-0.8205</td>
<td>0.0519</td>
</tr>
<tr>
<td></td>
<td>(0.1993)</td>
<td>(0.2129)</td>
</tr>
</tbody>
</table>

The test is invariant to the normalization chosen. Standard errors in parenthesis. The asterisks (***) indicate rejection of the null hypothesis of a unitary cointegrating parameter at the 1% significance level. Critical values (from the $\chi^2$ distribution): 3.8415 (5%) and 6.6349 (1%).

The bottom rows of Table 4.7 present the coefficients of the error correction terms (ECT) in both equations of the VAR. In the one case when these coefficients are significant, they display the correct sign. Since all variables are now I(0), statistical inference using standard $t$-tests is valid. The two countries display a completely opposite behaviour. While in Germany the government spending reacts strongly to budget imbalances, in the Netherlands it is the revenue side which assumes the burden of adjusting to equilibrate the government accounts, while the expenditures remain rather unresponsive. This analysis is carried further in the next section.

4.5. Causality effects in the relation revenues-expenditures

The causality tests in this section serve a double, interrelated purpose.\(^\text{12}\) Firstly, they are useful to confirm some of the econometric results of previous sections. The verification of no causal effects between the series of total revenues and total expenditures is an indication that these series do not have a long-run equilibrium relationship. From the

---

\(^{12}\) In what follows, the term causality is generally used in the econometric sense of 'Granger-causality'.

113
Engle and Granger (1987) representation theorem, if two series are cointegrated they can be modelled as an error correction model, which suggests that, at the very least, the lagged value of one variable must enter the other determining equation. The detection of no causal relations between two series found to be cointegrated in the previous section, would raise some doubts on the validity of the test results.

The second main purpose of these tests is to contribute new empirical evidence to the recent normative and positive discussions on the chronological ordering of the tax/spend decisions of the government. This question has usually been examined assuming that the IBC holds, i.e., only in countries with sustainable fiscal policies. Bohn (1991a), Koren and Stiassny (1995), Garcia and Hénin (1998), Payne (1998) and Ross and Payne (1998), for example, investigate this question only in those countries where the series of revenues and expenditures cointegrate. The objective is to assess how the governments usually react to fiscal imbalances, and therefore what policies should the economic agents expect in the future, following an increase in the deficit. Four alternative situations may emerge:

i) the government collects its revenues and then adjusts the expenditures so that the IBC is satisfied. The process of international economic integration may force this type of policy by restricting the autonomy of national policies, namely the capacity of the authorities to raise taxes above the levels of other countries. This hypothesis is known as the ‘tax-and-spend’ strategy, and it was first advanced by Friedman (1978) to explain the functioning of the fiscal process. He argues that an increase in taxes leads to an increase in spending and hence does not necessarily reduce the deficit;

ii) the government fixes the expenditures at a certain level and then adjusts the revenues according to the IBC. This is the ‘spend-and-tax’ strategy, advanced by Peacock and Wiseman (1979). They suggest that temporary increases in spending, following ‘crisis’, generate permanent increases in revenues, since expenditures do not return to their initial values. Barro’s (1979) tax smoothing model also implies a causality effect from expenditures to revenues;

iii) the decisions are mutually interdependent and are jointly determined. This is the view of the traditional public finance theory. Musgrave (1966), for example, argues that revenues and expenditures should be simultaneously decided, resulting from a comparison between the marginal costs and advantages of the government services;
iv) the choices on the value of expenditures and revenues are taken independently from each other, and therefore no causal relationship is to be expected.

This question has received large attention in the empirical literature since the mid-eighties, especially in the United States and Canada. Several tests have been performed to discriminate between the above four possible situations, but the results are far from conclusive. For the US, for example, all four hypothesis have found its supporters.\(^{13}\)

One other important implication of these tests concerns those countries intending to escape from an unsustainable fiscal policy, a perspective usually not considered in the literature. The success of a policy of fiscal consolidation based on increasing revenues or reducing spending may depend largely on the causal relationship between the variables. For example, the existence of a unidirectional positive causality effect running from revenues to expenditures implies that a policy of raising taxes may not be successful to curtail the budget, since it induces an increase in expenditures, eventually even more than offsetting. In this situation, it would be preferable to cut the government spending. Conversely, if it is the expenditures that cause revenues, an attempt to reduce the deficit by lower spending may be useless if revenues also decline.\(^{14}\)

Table 4.8 presents the results of the non-causality tests, following the methodology outlined in section 3.1.2. In the cases where the null hypothesis of non-cointegration was rejected, the tests are performed in the context of a VECM, which allows both short-term and long-term causality effects. For all the countries, the results of the LA-VAR test, introduced in section 3.1.2, were also reported. These tests were performed in a model with a constant term and no trend, but the results are robust to alternative specifications of the deterministic regressors. Including a time trend in the model, for example, did not change neither the lag structure nor the conclusions of the tests.\(^{15}\)


\(^{14}\) There is a possibility that a policy change may be 'structural', in the sense that it may also change the causal ordering. It is therefore important to avoid a merely mechanical interpretation of the results, and look at possible reasons for causal ordering.

\(^{15}\) The standard Granger-causality tests with the variables in first-differences were also performed, for the countries where no cointegrating relation was found, but the results were not included in the table for two
Table 4.8: Tests of non-causality between revenues and expenditures

<table>
<thead>
<tr>
<th>VECM</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>tg $\rightarrow$ t</td>
<td>n.c.¹</td>
<td>n.c.</td>
<td>1.3401</td>
<td>n.c.</td>
<td>15.5528 * * *</td>
<td>n.c.</td>
</tr>
<tr>
<td></td>
<td>[.247]</td>
<td></td>
<td></td>
<td></td>
<td>[.000]</td>
<td></td>
</tr>
<tr>
<td>t $\rightarrow$ tg</td>
<td>16.9423 * * *</td>
<td>0.7164</td>
<td>[.000]</td>
<td>[.699]</td>
<td>11.3401</td>
<td>[.247]</td>
</tr>
<tr>
<td>order</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LA-VAR</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>tg $\rightarrow$ t</td>
<td>1.3187</td>
<td>0.0666</td>
<td>0.3473</td>
<td>0.2927</td>
<td>4.8249 *</td>
<td>0.2890</td>
</tr>
<tr>
<td>t $\rightarrow$ tg</td>
<td>0.3246</td>
<td>0.1695</td>
<td>10.1498 * * *</td>
<td>0.5651</td>
<td>0.2101</td>
<td>0.0245</td>
</tr>
<tr>
<td>order²</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

¹ According to the Adjusted LR test (the SBC suggests order 2 for all countries, while the AIC indicates order 3 in some cases, without changing qualitatively the conclusions. ² n.c. indicates 'not cointegrated. The asterisks (*** and *) represent rejections of the null hypothesis of non-causality at the 10% and 1% levels of significance, respectively. The p-values are in brackets.

The results for the countries with an unsustainable fiscal policy do not identify any causal relationship between the variables. However, after introducing dummy variables in the model for Italy (in 1975 and 1982), in order to solve problems of non-normality, it is possible to identify a unidirectional causal relation running from revenues to expenditures.¹⁶ The LA-VAR test in this case produces a statistic of 6.2966, which rejects the null hypothesis of non-causality at the 1.2% significance level. This procedure of including dummies is not necessary in any other case, since no other problems of non-normality were detected in the residuals.

This result suggests that any policy of fiscal consolidation in Italy centred exclusively on raising taxes or seigniorage, for example, may not produce successful results. This increase in revenues is immediately followed by higher expenditures, expanding the share of the public sector in the economy, possibly without a deficit reduction.¹⁷ The

main reasons. First, all the model selection criteria chose order zero in all cases, immediately suggesting the non-existence of any causal relationship. Second, even disregarding the selection criteria and performing the tests with several different orders of the VAR did not produce any significant result.

¹⁶ Similar results for Italy were found by Owoye (1995), for example.

¹⁷ In fact, Tanzi and Fanizza (1996: p. 233) report that "Italy has been attempting to reduce its fiscal deficit mainly by increasing its level of taxation, which has risen sharply since 1980. Nonetheless, its fiscal deficit has remained high".
appropriate policy should involve a reduction in spending, as suggested for example in Alesina and Perotti (1997).

For the other three countries in this group, the causality tests do not provide any indication of the correct course of fiscal policy actions. The results suggest that the choice concerning the strategy for fiscal consolidation is neutral. However, they are useful to confirm the results of the cointegration tests in previous sections. The findings of non-causality in Belgium, France and the UK provide a supplementary indication that revenues and expenditures are not cointegrated in these countries. For the two latter countries, this removes some possible remaining doubts arising from the contradictory results of the Engle-Granger and the Johansen cointegration procedures.

In what concerns the two countries with sustainable fiscal policies, the tests based on the VECM representation in Table 4.8 produce strong evidence of a unidirectional causality relation. Identical results were obtained using the LA-VAR approach. The weaker rejection of the null hypothesis of non-causality in the case of the Netherlands may result, as noted before, from the lower power of this approach (see, for example, Mills, 1998, and Yamada and Toda, 1998). 18

This provides important indications in terms of both objectives set up in the beginning of this section. On the one hand, the existence of, at least unidirectional, causality is coherent with the results of cointegration between revenues and expenditures in Germany and the Netherlands. On the other hand, these results reveal a completely opposite adjustment policy in the two countries. In Germany, budgetary disequilibriums during this period have been corrected primarily by a reduction in government spending. 19 Conversely, in the Netherlands the burden of adjustment falls predominantly on the revenue side. 20 This may be one of the factors contributing to the share of the central government sector in the economy being significantly higher in the Netherlands than in Germany. In fact, during most part of the sample period, the fiscal policy in the Netherlands was basically directed at deficit targets, achieved mainly through tax

18 To assess the robustness of the results, other orders of the VECM were equally tested, using the cointegration vectors from Table 4.7 as the error correction term, and the 'seemingly unrelated regression' (SURE) approach for the estimation. In all cases, the results of Table 4.8 for Germany and the Netherlands were confirmed.
19 Ghosh (1995) interprets this causality relation, assuming the IBC holds, as evidence of tax-smoothing.
20 Similar results for the Netherlands were found by Joulaian and Mookerjee (1990).
increases (OECD Economic Surveys: Netherlands 1998, p. 48). The result was a rapid increase of the tax burden on the economy, evident in Figure A.1 of the appendix.

4.6. Alternative data sets

This section examines the robustness of the previous results to a change in the level of the government accounts (section 4.6.1) and in the frequency of the data (section 4.6.2). As before, the analysis follows the sequential testing procedure proposed in the previous chapter to evaluate the sustainability of the fiscal policies. Besides this purpose of 'confirmatory analysis', another important objective is to explore the additional information potentially present in these two data sets.

4.6.1. A broader level of government accounts: general government

The analysis below is focused on general government data, as described in section 3.2.2, which includes the financial balances of the central, regional and local governments, plus the social security funds. The main objective is to compare the results of the sustainability analysis in the two government levels, and the relative influence of the social security funds and the local governments' budgets on these results.

Figure A.4, in the appendix, displays the behaviour of the fiscal variables throughout the sample period available (starting in 1953 for Belgium, 1951 for Italy, and 1950 for all the other countries). The general government's total revenues and expenditures, as a share of GDP, increased continuously from the early fifties until the mid eighties. During the last decade they have stabilised at a value close to 50% of GDP in all countries except the UK (40%). A comparison with the more moderate evolution of the same variables with central government data (Figure A.1) reveals the increasing relative significance of the social security funds and the local governments in all these countries.
Figure 4.1 exhibits the relative importance of each main component of the general government finances, measured at the end of the sample period. The relative share of each component is obtained by adding the respective revenues and expenditures, and dividing the sum by the total value of the general government.

Figure 4.1: The main components of the general government’s budget in 1996

In Belgium, Italy, and especially the UK, the central government still administers more than half of the general government’s resources and applications. In Germany it manages a mere 20%, with the social security and the local governments sharing the rest of the budget in almost equal parts. This justifies the importance of this complementary analysis with general government data, and explains why the EU authorities insisted on the general government accounts being the basis of the fiscal convergence criteria.

As shown in Figure A.4, the growth rate of the expenditures has been on average slightly above the growth of revenues, resulting inevitably in an increasing deficit (Figure A.5). However, with the exception of Belgium, the average value of the general government’s deficit throughout the whole period was below the average value registered for the central government (Table 4.9).

---

21 The values for the central government were obtained by deducting the social security accounts and the
Table 4.9: Average deficit of the general and central governments in the period 1950-96

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>5.42</td>
<td>1.35</td>
<td>0.28</td>
<td>6.59</td>
<td>1.51</td>
<td>1.79</td>
</tr>
<tr>
<td>Central</td>
<td>5.16</td>
<td>2.22</td>
<td>1.15</td>
<td>7.23</td>
<td>1.92</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Values in per cent of GDP

This comparison suggests that, on average during this period, the combined budgets of the social security funds and the local governments have been on surplus. The central government is the section which is mainly responsible for the fiscal imbalances in these countries. Unfortunately, due to the unavailability of disaggregated data, it is not possible to examine the separate evolution of the deficits of each component during the whole sample. However, it is possible to observe that the situation is very different at the end of the sample period. Figure 4.2 presents the contribution of each component of the general government to the total deficit in 1996.

Figure 4.2: Total surplus of each component of the general government in 1996.


With the exception of the UK, the social security funds in all other countries were the main contributors to the deficits of the general government in 1996. The budgets of the regional and local governments are generally equilibrated, an observation which can be extended as far back as the early seventies, the earliest period when disaggregated data is extrabudgetary accounts from the consolidated central government.
available. Therefore, the tests in this section will essentially verify the influence of the social security funds in the sustainability analysis.

Following the testing strategy delineated in Figure 3.1 of the previous chapter, Table 4.10 presents the unit root tests on the series of total expenditures and revenues of the general government, using an ADF test with the lag length being chosen by the sequential procedure, ADF_{ST}. The trend variable is statistically significant in some of the regressions and therefore the table presents the results for both models with and without trend.

Table 4.10: Unit root tests on the revenues and expenditures series

<table>
<thead>
<tr>
<th>ADF_{ST}</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>total revenues t_{G}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no trend</td>
<td>-1.7819</td>
<td>-0.3912</td>
<td>-1.8210</td>
<td>-0.0849</td>
<td>-0.9939</td>
<td>-1.1679</td>
</tr>
<tr>
<td>trend</td>
<td>-0.1638</td>
<td>-2.0611</td>
<td>-1.5398</td>
<td>-1.2165</td>
<td>-0.4249</td>
<td>-2.1974</td>
</tr>
<tr>
<td></td>
<td>(0)\textsuperscript{1}</td>
<td>(0) \textsuperscript{1}</td>
<td>(0) \textsuperscript{1}</td>
<td>(0) \textsuperscript{1}</td>
<td>(0) \textsuperscript{1}</td>
<td>(1) \textsuperscript{1}</td>
</tr>
<tr>
<td>total expenditures t_{G}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no trend</td>
<td>-1.9565</td>
<td>-0.2570</td>
<td>-1.0072</td>
<td>-0.4561</td>
<td>-1.6811</td>
<td>-1.0981</td>
</tr>
<tr>
<td>trend</td>
<td>0.0525</td>
<td>-2.4901</td>
<td>-2.1456</td>
<td>-2.5923</td>
<td>-0.3115</td>
<td>-2.8233 \textsuperscript{*}</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(0)</td>
<td>(1)</td>
<td>(3)</td>
<td>(1)</td>
<td>(3)</td>
</tr>
</tbody>
</table>

\textsuperscript{1} The trend is not significant in the regression. The asterisk (*) indicates rejection of the null hypothesis of a unit root at the 10% significance level. Critical values (MacKinnon, 1991): -4.1695/-3.5778 (1%), -3.5134/-2.9256 (5%) and -3.1914/-2.6005 (10%) for the models with/without trend

It is not possible to reject the null hypothesis of a unit root in any of the series analysed in the table, whether or not a trend is included in the model.\textsuperscript{22} This conclusion is robust to alternative methods of choosing the lag length, such as for example the SBC. The second step in the testing strategy is to examine the two nested hypotheses of stationarity and stability of the debt ratio, by performing a unit root test on the series of total surplus of the general government. The results are displayed in Table 4.11, again for both models with and without trend.

---

\textsuperscript{22} The null hypothesis of a unit root in the first-differenced series was strongly rejected in all cases.
Table 4.11: Unit root tests on the total surplus series

<table>
<thead>
<tr>
<th>ADF$_{Tr}$</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>total surplus $t_s^G$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no trend</td>
<td>-1.5385</td>
<td>0.0362</td>
<td>-1.9929</td>
<td>-1.4768</td>
<td>-4.8338***</td>
<td>-2.8556*</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(1)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(1)</td>
</tr>
<tr>
<td>trend</td>
<td>-1.4878</td>
<td>-2.4782</td>
<td>-4.1366***</td>
<td>-1.6216</td>
<td>-4.1481***</td>
<td>-3.9904**</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(1)</td>
<td>(0)</td>
<td>(0)</td>
<td>(1)</td>
<td></td>
</tr>
</tbody>
</table>

The trend is not significant in the regression. In parenthesis is the number of lags. The asterisks (*** and **) indicate rejection of the null hypothesis of a unit root at the 1% and 5% significance levels, respectively. Critical values (MacKinnon, 1991): -4.1695/-3.5778 (1%), -3.5134/-2.9256 (5%) and -3.1914/-2.6005 (10%) for the models with/without trend.

According to the results in this table, the series of total surplus for the Netherlands is clearly stationary, whatever the model used in the estimation, while for Germany and the UK, the hypothesis of a unit root is only rejected when a trend is included in the model.

Table 4.12 investigates whether the presence of a structural break in the model changes the conclusions of the unit root tests.

Table 4.12: Unit root tests allowing a structural break in the series

<table>
<thead>
<tr>
<th>Additive Outlier Model (AO)</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test statistic $t_{SE}$</td>
<td>-2.5956</td>
<td>-3.3544</td>
<td>-4.2607</td>
<td>-2.8794</td>
<td>-5.2498***</td>
<td>-4.6631*</td>
</tr>
<tr>
<td>(lag length $p$)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(3)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>Dummy coef. $\beta$</td>
<td>-0.0374</td>
<td>-0.0236</td>
<td>-0.0540</td>
<td>-0.0703</td>
<td>-0.0458</td>
<td>-0.0237</td>
</tr>
<tr>
<td>(stand. error)</td>
<td>(0.0061)</td>
<td>(0.0040)</td>
<td>(0.0070)</td>
<td>(0.0078)</td>
<td>(0.0084)</td>
<td>(0.0065)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Innovative Outlier Model (IO)</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test statistic $t_{SE}$</td>
<td>-2.9315</td>
<td>-3.3242</td>
<td>-4.1827</td>
<td>-4.3353</td>
<td>-5.5157***</td>
<td>-4.4251*</td>
</tr>
<tr>
<td>(lag length $p$)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td></td>
</tr>
<tr>
<td>Dummy coef. $\beta$</td>
<td>-0.0117</td>
<td>-0.0165</td>
<td>-0.0319</td>
<td>-0.0380</td>
<td>-0.0109</td>
<td>-0.0120</td>
</tr>
<tr>
<td>(stand. error)</td>
<td>(0.0061)</td>
<td>(0.0059)</td>
<td>(0.0088)</td>
<td>(0.0098)</td>
<td>(0.0044)</td>
<td>(0.0037)</td>
</tr>
</tbody>
</table>

The AO (IO) model assumes an instantaneous (gradual) change in the level of the series. The number of lags was chosen so that the coefficient on the last included lag is significant at the 10% level of significance (starting from a maximum of three). The asterisks (*** indicate rejection of the null hypothesis of a unit root at the 1% significance level. Critical values (Perron and Vogelsang, 1992a): -5.20/-5.51 (1%), -4.67/-4.76 (5%) and -4.33/-4.42 (10%) for the AO/IO models.
Both AO and IO models present similar results, with a strong rejection of the unit root null only in the case of the Netherlands. This indicates that only this country clearly satisfies both sustainability conditions. Further inferences may result from the cointegration tests in the next tables, which allow the cointegrating coefficients to be freely determined by the regression. The subsequent tests on these coefficients provide an alternative joint test of the two sustainability conditions.  

Table 4.13: Engle-Granger cointegration tests

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tt = \alpha t + \beta tt$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF test</td>
<td>-2.0749</td>
<td>-2.6266</td>
<td>-4.8637***</td>
<td>-1.2674</td>
<td>-4.1083**</td>
<td>-3.4971***</td>
</tr>
<tr>
<td>parameters:</td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
</tr>
<tr>
<td></td>
<td>0.0321</td>
<td>0.0596</td>
<td>0.1539</td>
<td>0.0739</td>
<td>0.1015</td>
<td>0.1099</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0108)</td>
<td>(0.0104)</td>
<td>(0.0123)</td>
<td>(0.0130)</td>
<td>(0.0189)</td>
</tr>
<tr>
<td></td>
<td>0.8159</td>
<td>0.8284</td>
<td>0.6164</td>
<td>0.6457</td>
<td>0.7363</td>
<td>0.6618</td>
</tr>
<tr>
<td></td>
<td>(0.0230)</td>
<td>(0.0251)</td>
<td>(0.0252)</td>
<td>(0.0302)</td>
<td>(0.0284)</td>
<td>(0.0495)</td>
</tr>
<tr>
<td>$tg = \alpha t + \beta tt$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF test</td>
<td>-2.0621</td>
<td>-2.6056</td>
<td>-4.6765***</td>
<td>-1.4853</td>
<td>-4.7721***</td>
<td>-3.4343*</td>
</tr>
<tr>
<td>parameters:</td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
</tr>
<tr>
<td></td>
<td>-0.0229</td>
<td>-0.0522</td>
<td>-0.2037</td>
<td>-0.0699</td>
<td>-0.1014</td>
<td>-0.0566</td>
</tr>
<tr>
<td></td>
<td>(0.0142)</td>
<td>(0.0147)</td>
<td>(0.0251)</td>
<td>(0.0222)</td>
<td>(0.0214)</td>
<td>(0.0327)</td>
</tr>
<tr>
<td></td>
<td>1.1859</td>
<td>1.1592</td>
<td>1.5092</td>
<td>1.4132</td>
<td>1.2728</td>
<td>1.2069</td>
</tr>
<tr>
<td></td>
<td>(0.0335)</td>
<td>(0.0351)</td>
<td>(0.0616)</td>
<td>(0.0660)</td>
<td>(0.0491)</td>
<td>(0.0903)</td>
</tr>
</tbody>
</table>

The number of lags used in the ADF test (in parenthesis) is suggested by the SBC. Standard errors, in parenthesis, below the parameters’ estimates. The asterisks (***), (**), and (*) indicate rejection of the null hypothesis of no-cointegration at, respectively, the 1%, 5% and 10% significance levels. Critical values (MacKinnon, 1991): -4.1433 (1%), -3.4717 (5%) and -3.1374 (10%).

In Germany and the Netherlands, the Engle-Granger two-step testing procedure clearly suggests the existence of a cointegrating relationship between the series of revenues and expenditures. The null hypothesis of a unit root in the residuals of the cointegrating equation is also rejected in the case of the UK, but the statistics are too close to the

---

23 From the Johansen procedure, since it is not valid to carry out inference on the cointegrating coefficients of the Engle-Granger OLS regression, using the usual t-tests.
critical values to be conclusive. As happened before with central government data, further tests are necessary to allow a definitive conclusion in this latter case.

The Johansen test in Table 4.14 may help reach a final decision. The upper part of the table presents the results of the maximum eigenvalue test of the null hypothesis of zero against the alternative of one cointegrating vector. The Trace-test delivered basically the same conclusions (the only difference being that in model $H_I^*(r)$ for Italy, the Trace-test rejects the null at only the 10% level of significance). The tests of the null hypothesis of one or less cointegrating relations against the alternative of two vectors were not rejected in any case and therefore have also been omitted from the summary Table 4.14, for the sake of brevity.

Given the possibility that some of the variables may be trended, following the results of the unit root tests, it is necessary to devote special attention to the choice of deterministic terms to be included in the model. The tests are performed with three different specifications of the deterministic terms. Following the notation in Johansen, $H^*(r)$ is the model with an unrestricted intercept and a trend in the cointegrating vector, $H_I(r)$ has an unrestricted intercept but no trend, and $H_I^*(r)$ displays a restricted intercept in the cointegrating vector.

The problem is immediately solved if all specifications deliver similar conclusions. However, in two cases, the final conclusions of the tests depend on the underlying model considered. Therefore, to choose the correct model, Table 4.14 also presents some tests to choose the correct model, following the 'Schwarz Bayesian Criterion', and the likelihood ratio procedure suggested in Johansen (1994). This latter procedure was described in section 3.1.2.

The only contradictory results appear in Belgium and Italy, where the null hypothesis of no cointegration is not rejected in models $H^*(r)$ and $H_I(r)$ but is (marginally) rejected with model $H_I^*(r)$. However, given the choices of the SBC and the likelihood ratio tests, it seems reasonable to conclude against the existence of a cointegrating relationship in any of these two countries.
Table 4.14: Johansen’s maximum eigenvalue cointegration test (λ_{MAX} test)

<table>
<thead>
<tr>
<th>model</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>H'(r)</td>
<td>10.4216</td>
<td>7.0920</td>
<td>22.8448''</td>
<td>10.3323</td>
<td>17.0430</td>
<td>15.1571</td>
</tr>
<tr>
<td>H(r)</td>
<td>5.1675</td>
<td>6.8131</td>
<td>22.4748''</td>
<td>4.8254</td>
<td>15.1889''</td>
<td>11.8161</td>
</tr>
<tr>
<td>H_1(r)</td>
<td>16.5065''</td>
<td>11.9213</td>
<td>24.3519''</td>
<td>17.4623''</td>
<td>15.7372'</td>
<td>11.8167</td>
</tr>
</tbody>
</table>

Schwarz Bayesian Criterion

| H'(r)          | 257.8245      | 299.3713      | 277.1180      | 265.8059      | 253.2897     | 263.4388  |
| H(r)           | 257.8245      | 299.3713      | 278.8473      | 265.8059      | 254.2660     | 263.6368  |
| H_1(r)         | 255.7949      | 297.4214      | 277.9355      | 263.8819      | 255.7433     | 267.1308  |

Likelihood ratio procedure

| H'(r)/H(r)      | 5.2544''      | 0.2793        | 0.3696        | 5.5068''      | 0.8521       | 3.3410    |
| H_1(r)/H_1(r)   | 1.8232        | 6.4489''      | 5.6522''      | 1.0837        | 1.8540       | 0.2287    |

The asterisks (**) and (*) in the upper part of the table indicate rejection of the null hypothesis of no cointegration at the 5% and 10% levels of significance, respectively. Critical levels (Pesaran, Shin and Smith, 1997): 19.2211/4.8811/5.87 (5%) and 17.18/12.98/13.81 (10%) for the λ_{MAX} test in models H'(r)/H(r)/H_1(r), respectively. The asterisks (**) in the lower part of the table indicate rejection of the null hypothesis that the restricted model on the right is more adequate, at the 5% significance level. Critical values (from the \chi^2 distribution): 6.6349 (1%) and 3.8415 (5%).

In Germany and the Netherlands, there is evidence of a cointegrating relationship, confirming the conclusions of sustainability obtained with the central government data set. For these two countries, Table 4.15 presents the Wald tests on the parameters of the cointegrating vector to verify whether the supplementary condition of sustainability (boundness of the debt ratio) is also satisfied. These tests are performed in both the models with a restricted and an unrestricted intercept, and no trend, following the indications of the previous table.24

The results in this table are completely opposite to the ones obtained with central government data in Table 4.7. The hypothesis of a unitary cointegrating coefficient is strongly rejected for Germany but not for the Netherlands. This confirms the results of the unit root tests on the series of the total surplus (Table 4.11), which strongly rejected the hypothesis of nonstationarity for the Netherlands, while for Germany the result was dependent on the specification of the model. Moreover, the results are fully compatible

---

24 The trend variable in the cointegrating vector is not significant in both cases and, besides, using model H'(r) yields qualitatively the same results in both countries.
with the outcome of the unit root tests when a structural break is allowed in the model (Table 4.12).

Table 4.15: Tests on the parameters of the error correction terms

<table>
<thead>
<tr>
<th>Parameters:</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>LR test of restrictions</th>
<th>ECT coefficient in:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-tt + \alpha + \beta tt) (^1)</td>
<td>( H_1(r) )</td>
<td>( H_1^*(r) )</td>
<td>( H_1(r) )</td>
<td>( H_1^*(r) )</td>
</tr>
<tr>
<td>Germany</td>
<td>0.1646 (0.0150)</td>
<td>0.6016 (0.0346)</td>
<td>0.8569 (0.0519)</td>
<td>0.3530 (0.0871)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.0378 (0.0150)</td>
<td>-0.2225 (0.1137)</td>
<td>-0.2401 (0.1098)</td>
<td>-0.2401 (0.1098)</td>
</tr>
</tbody>
</table>

\(^1\) The test is invariant to the normalisation chosen. Standard errors in parenthesis. The asterisks (***) and (*) indicate rejection of the null hypothesis of a unitary cointegrating parameter at the 1% and 10% significance levels, respectively. Critical values (from the \( \chi^2(1) \) distribution): 3.8415 (5%) and 6.6349 (1%).

This inversion in the results may be caused by the inclusion of the social security and the local government financial balances in the general government definition. According to the analysis in the beginning of this section, it seems more reasonable to point at the social security funds as the origin of the problems in Germany.\(^{25}\) On the other hand, the indication of a better fiscal situation in the Netherlands with general than with central government data, probably reflects the important structural reforms in the social security system since 1982. These reforms were integrated in a major change in economic policy towards fiscal consolidation of the whole public sector (\textit{OECD Economic Surveys: Netherlands, 1998}), and are clearly discernible in Figure A.4 in the appendix.

\(^{25}\) "Germany's welfare timebomb" (The Economist, 5th June 1999).
The last rows in Table 4.15 also suggest a change in the way both countries react to a budgetary disequilibrium. This can be confirmed in Table 4.16, where the causality tests in the context of the VECM and the LA-VAR models are presented.26

Table 4.16: Tests of non-causality between revenues and expenditures

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>VECM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$tg \rightarrow u$</td>
<td>n.c.</td>
<td>n.c.</td>
<td>16.4247***</td>
<td>n.c.</td>
<td>7.4883**</td>
<td>n.c.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[.000]</td>
<td></td>
<td>[.024]</td>
<td></td>
</tr>
<tr>
<td>$u \rightarrow tg$</td>
<td>5.0423**</td>
<td>4.7814*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[.025]</td>
<td></td>
<td>[.092]</td>
<td></td>
</tr>
<tr>
<td>order</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA-VAR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$tg \rightarrow u$</td>
<td>2.5538</td>
<td>4.9388**</td>
<td>5.2344**</td>
<td>0.8620</td>
<td>7.6111**</td>
<td>0.5836</td>
</tr>
<tr>
<td>$u \rightarrow tg$</td>
<td>0.2651</td>
<td>0.0657</td>
<td>0.5248</td>
<td>0.9940</td>
<td>3.9894</td>
<td>1.2363</td>
</tr>
<tr>
<td>order$^*$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

$^*$ According to the Adjusted LR test. $^1$ n.c. indicates 'not cointegrated'. The asterisks (***), (**) and (*) represent rejections of the null hypothesis of non-causality at the 1%, 5% and 10% levels of significance, respectively. The p-values from the $\chi^2$ distribution are in square brackets.

The results of the non-causality tests reveal that for France, Germany and the Netherlands, there is a significant causal relationship running from expenditures to revenues. In these two last countries, there is also some evidence of the inverse causal relation, but the rejection is less powerful and occurs only with the VECM. A comparison with the results using central government data reveals again some differences. While for France the earlier tests did not disclose any causal relation, for Germany the main direction of causality has been reversed.

These results may be explained by the lower flexibility of the social security and local governments expenditures. The adjustment burden in this situation falls primarily on the revenue side. An unfunded social security system, for example, implies a 'pay as you go' scheme, i.e., get the revenues necessary to respond to the spending needs.

26 According to the model selection criterion in Table 4.14, model $H_4(r)$ was used for Germany and model $H_{41}(r)$ for the Netherlands.
In sum, this section has shown that when general government data is considered, Germany and the Netherlands continue to be the only two countries in the sample with clear indications of a sustainable fiscal policy. However, the tests have also revealed that Germany may face future problems of unbounded growth of its debt-GDP ratio, while the Netherlands seems to satisfy both sustainability criteria.

This establishes an important finding of the section. This reversal of fiscal positions, comparing with the results obtained with central government data, underlines the relatively high importance of the social security funds on the financial balances of the general government. In particular, it was shown above that, recently, the social security was the main contributor to the deficits of the general government in most countries.

Another important difference relatively to the tests with the main data set, particularly noted in the case of Germany, is that now the burden of the fiscal adjustment falls predominantly on the revenue side. This suggests a lower flexibility of the general government's expenditures, and may help explain the significant increase in the share of this sector in the economic activity during the last decades.

4.6.2. A larger number of observations: quarterly data

The quarterly series of total revenues and total expenditures of the central government, as ratios to GDP, are displayed in figure A.6 of the appendix for all six EU countries. The quarterly fiscal variables were taken from the same source of the annual series, and therefore display the same long-term behaviour in the period 1957-96. The main interest of these graphs is therefore to reveal the strong seasonal patterns in the quarterly variables, across the whole sample.

4.6.2.1. Seasonality issues

To avoid the interference of these seasonal effects in the long-term analysis, the usual procedure is to remove them by pre-adjusting the data with some filter. For example, using a ratio to moving-average adjustment methodology, yields the adjusted series in
Both series were divided by GDP, the only variable which comes already adjusted from the source. Applying the ADF\textsubscript{ST} test to these two series, and also to the difference between them, the total surplus, produces the results shown in Table 4.17. This sequential test begins with a maximum of 12 lags, and uses a model without a trend variable, since this was not significant in the regressions.

<table>
<thead>
<tr>
<th>Table 4.17: ADF\textsubscript{ST} unit root tests on total revenues and expenditures, seasonally adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
</tr>
<tr>
<td>revenues ( r_t )</td>
</tr>
<tr>
<td>(7)</td>
</tr>
<tr>
<td>expenditures ( g_t )</td>
</tr>
<tr>
<td>(7)</td>
</tr>
<tr>
<td>surplus ( s_t )</td>
</tr>
<tr>
<td>(7)</td>
</tr>
</tbody>
</table>

The ADF\textsubscript{ST} uses a sequential procedure to choose the lag length, starting from a maximum of 12. The asterisk (*) indicates rejection of the null hypothesis of a unit root at the 10% significance level. Critical values for 159 values (MacKinnon, 1991): -3.4640 (1%), -2.8748 (5%) and -2.5749 (10%).

These results are very similar to the ones previously obtained with annual data. The main difference is that the hypothesis of a unit root is not rejected in any series, at the usual levels of significance, while before the nonstationarity of the surplus in Germany was strongly rejected. Major differences, however, were found by the cointegration tests in Table 4.18. The table presents only the results for the maximal eingenvalue test of the null hypothesis of no cointegration, for the sake of brevity. The Trace test delivered exactly the same results and the null hypothesis of one or less cointegrating vectors against the alternative of more than one vector was never rejected. The model includes a restricted intercept and no time trend.

27 The data were adjusted with the seasonal adjustment procedure Samas of TSP, also used for the same purpose in Haug (1995). Alternatively, the adjustment routine Esmooth in the package Rats was also employed, but did not produce significant differences in the results of the tests.

28 Adjusting the already adjusted series of GDP leaves the series practically unchanged (the 'adjustment factors' are all very close to unity). Repeating the tests using this re-adjusted GDP series does not alter any of the conclusions.

29 Applying the ADF\textsubscript{STC} procedure yields qualitatively equivalent results, albeit in more parsimonious models.
Table 4.18: Johansen cointegration tests, seasonally adjusted data

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>data seasonally adjusted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{\text{MAX}}$ test</td>
<td>5.0184</td>
<td>16.3999&quot;</td>
<td>15.7731&quot;</td>
<td>10.2333</td>
<td>12.9989</td>
<td>13.8370&quot;</td>
</tr>
<tr>
<td>including seasonal dummies with unadjusted data</td>
<td>(6)</td>
<td>(4)</td>
<td>(5)</td>
<td>(4)</td>
<td>(5)</td>
<td>(5)</td>
</tr>
<tr>
<td>$\lambda_{\text{MAX}}$ test</td>
<td>17.9835&quot;</td>
<td>33.6316&quot;</td>
<td>38.2137&quot;</td>
<td>29.1866&quot;</td>
<td>32.4313&quot;</td>
<td>15.7005&quot;</td>
</tr>
<tr>
<td>including seasonal dummies</td>
<td>(4)</td>
<td>(4)</td>
<td>(4)</td>
<td>(4)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
</tbody>
</table>

In parenthesis is the order of the VAR. The asterisks (**) and (*) indicate rejection of the null hypothesis of no cointegration against the alternative of one cointegrating vector at the 5% and 10% significance levels. Critical values (Pesaran, Shin and Smith, 1997): 15.87 (5%) and 13.81 (10%).

The upper part of the table displays the cointegration tests on the seasonally adjusted series. Only in the case of France it is possible to reject the hypothesis of no cointegration at the usual levels of significance. This disparity in the results relatively to the annual data may be caused by the seasonal adjustment method. Comparing figures A.6 and A.7, it becomes apparent that the adjustment has not been efficient in removing the seasonal effects from the original series.

As argued by Franses (1996), for some variables it may be very difficult to disentangle seasonal and nonseasonal effects. In these situations, it may be more appropriate to explicitly incorporate seasonality in the econometric model (using seasonal dummies, for example) than to remove it with some pre-adjustment method. The lower part of Table 4.18 displays the results of the same cointegration tests in a model including seasonal dummies. The tests strongly suggest the existence of a cointegrating relationship in all countries, except the UK.

All these unexpected results may be the consequence of a changing seasonal pattern throughout the sample period. Seasonal adjustment using a moving-average filter or employing seasonal dummy variables presupposes that the seasonal pattern is constant or

---

30 If these changes are discrete and easily identifiable, one possible solution is to adjust the series for different separate periods. Other hypothesis is to pre-decompose the series. The seasonal pattern of these series may vary due to a change in the seasonal pattern of one or more components of the series, or due to a change on its relative composition, even considering that each component has a different but constant seasonal pattern. Considering these hypotheses, Joines (1990) adjusts the series of the deficit by decomposing it into several different components and different sub-periods, and adjusting each sub-period of each component in turn. Even disregarding the difficulty in dividing a long time series into several different components, this procedure is completely arbitrary in the choice of the components and the truncation periods, a decision which may have a non-negligible effect, distorting the outcome of the tests.
stationary over time. If this is not the case, the filter may change the stochastic properties of the series. It has been found, for example, that pre-adjusting the data with a moving-average process induces a reduction in the power of the unit root tests to reject the nonstationarity hypothesis (see, for example, Ghysels and Perron, 1993 and 1996, Canova and Hansen, 1995, and Olekalns, 1996). The evidence on cointegration tests has not yet been clearly analysed, but it seems plausible that two series adjusted with the same filter may appear to be cointegrated while in reality they are not (Wells, 1997).

Figure 4.3: Total revenues and expenditures in France (quarterly growth rates)

a) Total revenues

The lines represent the first differences of the total revenues and expenditures (in logs), in each quarter.

b) Total expenditures

To identify the type of seasonal pattern, a simple procedure was suggested by Franses (1994).\textsuperscript{31} It basically consists of isolating each quarter (in logs) and graphically displaying

\textsuperscript{31} For applications see, for example, Hylleberg (1992), Engle, Granger, Hylleberg and Lee (1993), Canova and Ghysels (1994), and Franses (1996).
the four annual series obtained, as well as their first-differences.\textsuperscript{32} As an example, Figure 4.3 presents this decomposition in the specific case of France, the only country where a cointegrating vector was found with seasonally adjusted data.

With a deterministic seasonal process, the lines would not cross and would maintain an almost constant amplitude between them. When the lines cross, as often happens in the graphs, it suggests a nonstationary seasonal process.\textsuperscript{33} Therefore, the tests based on the assumption of a deterministic seasonal pattern, as in Table 4.17 and Table 4.18, may produce spurious results and should be considered very cautiously.

4.6.2.2. Tests of sustainability with seasonal unit roots

A more formal procedure to analyse the seasonal patterns of a time series was first proposed by Hylleberg, Engle, Granger and Yoo (1990), henceforth HEGY. They extended the unit root testing methodology of Dickey and Fuller (1979) to the seasonal case, by factorizing the seasonal differencing polynomial \((1-L^4)=(1-L)(1+L)(1+L^2)\), where \(L\) is the backward shift operator.\textsuperscript{34} The first factor \((1-L)\) removes the unit root at frequency zero, leaving the seasonal roots unaffected. The factor \((1+L)\) corrects for the unit root at frequency \(\frac{1}{2}\) (semi-annual), not affecting the others. Finally, \((1+L^2)\) adjusts for the pair of complex conjugate unit roots at frequency \(\frac{1}{4}\) and \(\frac{3}{4}\) (annual). The HEGY tests are based on the regression

\[
y_{4t} = \alpha + \sum_{j=1}^{3} \beta_j D_j + \gamma_1 y_{1t-1} + \gamma_2 y_{2t-1} + \gamma_3 y_{3t-2} + \gamma_4 y_{4t-1} + \sum_{j=1}^{k} \delta_j y_{4t-j} + \epsilon_t,
\]

where \(y_i\) are auxiliary variables, \(y_1=(1+L+L^2+L^3)x_t\), \(y_2=-(1-L+L^2-L^3)x_t\), \(y_3=-(1-L^2)x_t\), \(y_4=(1-L^4)x_t\), \(x_t\) is the original series, and \(D_j\) represent seasonal dummies. The number of lagged dependent variables, \(k\), is chosen with the aim of whitening the residuals. This regression model can be estimated using ordinary least squares.

---

\textsuperscript{32} This graphical analysis does not require stationarity conditions, but the graph may be more easily interpreted after the long-run characteristics of the series are filtered out by first-differencing the log of the series (Canova and Ghysels, 1994).

\textsuperscript{33} The same phenomenon happens for all other countries and variables.

\textsuperscript{34} The assumption of a certain differencing filter, e.g. \((1-L^4)\) amounts to an assumption of the number of seasonal and nonseasonal unit roots in a time series (Franses, 1996).
The tests on the existence of a unit root at the zero (long-run) and semi-annual frequency are based on the ‘$t$ statistic’ of the estimates of $\gamma_1$ and $\gamma_2$, respectively, while the test of a unit root at the annual frequency is based on an ‘$F$ test’ of the joint significance of $\gamma_3$ and $\gamma_4$. The finite sample distributions for these $t$- and $F$-statistics are non-standard and are tabulated by Monte Carlo simulations in HEGY.

The existence of a seasonal unit root implies that the series has a seasonal pattern which changes over time, and certain shocks may cause these seasonal patterns to change permanently. These tests for seasonal roots allow the determination of the appropriate filter to use in the series, in contrast to the ad hoc filters used above. Table 4.19 presents the results of this test on the series of total revenues and total expenditures, as ratios to GDP. Following HEGY, Engle et al. (1993) and Franses (1996), the number of lagged dependent variables included in the regression was chosen recursively, starting from 12 lags and rejecting successively all non-significant lags (at the 10% level of significance) except those necessary to ensure no serial correlation and heteroscedasticity in the residuals.35

The null hypothesis of a unit root in the series of total revenues and expenditures cannot be rejected at the long-run for any country (first line in each variable), and at the seasonal frequencies for several countries. This result suggests that it is not correct to model these time series by including seasonal dummies in the regression, as if the seasonal process was deterministic. The usual tests of cointegration at the long-run must be applied only after the series are properly filtered, to remove the seasonal roots indicated by the tests.36 Only in the case of Germany, where the seasonal unit roots were all rejected, can the tests be applied to non-filtered data.

Before that, it is possible to test whether both sustainability conditions (cointegration and a unitary cointegrating vector) hold, by applying these same unit root tests to the series of total surplus. The results are presented in the lower part of Table 4.19, and show that, as

35 With an alternative strategy, starting from 12 lagged dependent variables and testing down sequentially until a significant lag is obtained, the results are only slightly different (the test $F$ for the variable $Btt$ and the test $t$ on $\gamma_3$ for the variables $Gtg$, $Ntt$ and $UKtg$, do not reject the hypothesis of non-significant parameters). It will be checked later whether these differences affect the results of the cointegration tests.

36 To test for ‘seasonal cointegration’ it would be necessary to filter the series to remove the possible unit root at frequency zero as well as at the seasonal frequency not being tested. The existence of seasonal
happened before with the annual central government data, only Germany seems to obey both conditions.

Table 4.19: HEGY unit root tests on the fiscal variables

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>total revenues $t_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-1.7497</td>
<td>-1.7259</td>
<td>-0.2904</td>
<td>-0.2226</td>
<td>-1.6320</td>
<td>-1.7145</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.6252</td>
<td>-2.5417</td>
<td>-3.6886**</td>
<td>-1.2670</td>
<td>-3.3263**</td>
<td>-1.0192</td>
</tr>
<tr>
<td>$\gamma_3\gamma_4$</td>
<td>8.7368**</td>
<td>6.5205*</td>
<td>12.3422***</td>
<td>5.6055*</td>
<td>5.2187</td>
<td>4.5998</td>
</tr>
<tr>
<td>lags</td>
<td>4</td>
<td>4.8</td>
<td>10</td>
<td>4</td>
<td>12,3,5,6</td>
<td>1,4,5,8</td>
</tr>
<tr>
<td>total expenditures $t_g$,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-1.2073</td>
<td>-2.1604</td>
<td>-1.1460</td>
<td>-0.9474</td>
<td>-1.4378</td>
<td>-1.7872</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-3.2321**</td>
<td>-1.1940</td>
<td>-4.4927***</td>
<td>-2.0296</td>
<td>-4.7792***</td>
<td>-5.4664***</td>
</tr>
<tr>
<td>$\gamma_3\gamma_4$</td>
<td>1.6222</td>
<td>-3.7732</td>
<td>14.2123***</td>
<td>4.3706</td>
<td>15.9289***</td>
<td>31.4941***</td>
</tr>
<tr>
<td>lags</td>
<td>2,3,4</td>
<td>1.2,4,5</td>
<td>2.5,9</td>
<td>1,2,4</td>
<td>1,4,5,8</td>
<td></td>
</tr>
<tr>
<td>total surplus $t_s$,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-1.2662</td>
<td>-1.9771</td>
<td>-3.1190**</td>
<td>-1.6701</td>
<td>-2.3144</td>
<td>-2.8036*</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.5082</td>
<td>-1.6576</td>
<td>-1.8525</td>
<td>-2.6849*</td>
<td>-3.8624***</td>
<td>-1.7321</td>
</tr>
<tr>
<td>$\gamma_3\gamma_4$</td>
<td>2.8163</td>
<td>2.5807</td>
<td>10.8202***</td>
<td>8.7014**</td>
<td>8.7478**</td>
<td>4.0156</td>
</tr>
<tr>
<td>lags</td>
<td>2-4,7,10</td>
<td>1,2,4</td>
<td>1,8,9</td>
<td>4,8,11</td>
<td>1</td>
<td>1,3,4,6</td>
</tr>
</tbody>
</table>

The series Ntg, UKtt and UKtg have problems of serial correlation or heteroscedasticity which only disappear with the inclusion of 'local' dummies (1995Q4 and 1996Q4 for Ntg; 1963Q1 and 1964Q1 for UKtt; 1963Q1 and 1963Q2 for UKtg) - however, the inclusion of these dummies does not affect the conclusions of the tests in any case. The asterisks (**), (*) indicate rejection of the null hypothesis of a unit root at the corresponding frequency at the 1%, 5% and 10% levels of significance, respectively. Critical values for $T=136$ (from HEGY): $-3.56/-3.49/8.92$ (1%), $-2.94/-2.90/6.63$ (5%) and $-2.62/-2.59/5.56$ for the tests on $\gamma_1/\gamma_2/\gamma_3\gamma_4$, respectively.

The non-rejection of unit roots at the long-run and seasonal frequencies may be a consequence of the neglected existence of structural breaks in the data (Smith and Otero, 1997). To account for these breaks, the analysis below follows the methodology suggested by Franses and Vogelsang (1998) to test for unit roots in quarterly data, considering the presence of one deterministic mean shift in one or more seasons. The cointegration between two time series indicates a varying but parallel seasonal pattern (see, for example, Engle et al., 1993).
analysis allows the unknown date of the break to be endogenously determined by the
data.\textsuperscript{37}

The test of the null hypothesis of unit roots at all frequencies depends on the underlying
model considered. In the case of the Additive Outlier (AO) model, where the shifts are
instantaneous, the testing procedure follows essentially two steps. The first consists of
removing the deterministic part from the series (of the total surplus, $t_s$, in this case) by
the auxiliary regression

$$
t_s = \sum_{s=1}^{4} \gamma_s D_s + \sum_{s=1}^{4} \delta_s DU_s + \tilde{y}
$$

and then use the residuals in the regression

$$
\Delta_s \tilde{y}_t = \pi_1 \tilde{y}_{t-1} + \pi_2 \tilde{y}_{t-2} + \pi_3 \tilde{y}_{t-3} + \pi_4 \tilde{y}_{t-4} +
$$

$$
+ \sum_{i=1}^{s} \phi_i D(TB)_s + \sum_{i=1}^{s} \eta_i D(TB)_{s+i} + \sum_{j=1}^{s} c_j \Delta_i \tilde{y}_{t-j} + e_t
$$

where $D_{s,t}$ are ordinary seasonal dummies, $DU_{s,t}$ are seasonal dummies which only take
nonzero values when $t>T_B$, $T_B$ is the date of the break, $\tilde{y}$ are the residuals from the first
regression, $D(TB)_{s,t}=\Delta_t Du_{s,t}$, and

$$
\tilde{y}_{t} = (1 + B + B^2 + B^3) \tilde{y}_t
$$

$$
\tilde{y}_{2,t} = -(1 - B + B^2 - B^3) \tilde{y}_t
$$

$$
\tilde{y}_{3,t} = -(1 - B^2) \tilde{y}_t
$$

The number of lagged dependent variables ($k$) is chosen following the sequential test
used before. Starting from a maximal value, $k_{MAX}=5$ as suggested in Franses and
Vogelsang (1998), $k$ is chosen so that the coefficient of the last lagged term is significant
at the 10\% significance level.

As in HEGY, the statistics of interest in regression (4.11) are $t$-statistics for the first two
coefficients $t_{m1}$, $t_{m2}$, and the joint statistic $F_{34}$. The upper half of Table 4.20 presents these
statistics, as well as the break-dates chosen endogenously by the tests. These break-dates

\textsuperscript{37} Franses and Vogelsang (1995) consider the tests for a known break-date.
are selected by minimising the $t$-statistics and maximising the $F$-statistic. The critical values are from Table B.1 in Franses and Vogelsang (1998).

Table 4.20: Long-run and seasonal unit root tests with a structural break

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive Outlier (AO) model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min $t_{cl}$</td>
<td>-2.2813</td>
<td>-3.6019</td>
<td>-4.9361**</td>
<td>-3.1748</td>
<td>-3.0303</td>
<td>-3.8067</td>
</tr>
<tr>
<td>[1972Q3]</td>
<td>[1979Q4]</td>
<td>[1974Q4]</td>
<td>[1968Q1]</td>
<td>[1972Q3]</td>
<td>[1973Q4]</td>
<td></td>
</tr>
<tr>
<td>Min $t_{c2}$</td>
<td>-2.6721</td>
<td>-3.6184</td>
<td>-5.6154**</td>
<td>-4.9461**</td>
<td>-4.6321**</td>
<td>-2.5981</td>
</tr>
<tr>
<td>[1973Q2]</td>
<td>[1972Q4]</td>
<td>[1978Q2]</td>
<td>[1978Q4]</td>
<td>[1979Q4]</td>
<td>[1973Q3]</td>
<td></td>
</tr>
<tr>
<td>Max $F_{cl}$</td>
<td>10.0649</td>
<td>7.7408</td>
<td>32.7297**</td>
<td>24.6338**</td>
<td>15.4805**</td>
<td>15.3173**</td>
</tr>
<tr>
<td>[1973Q1]</td>
<td>[1974Q3]</td>
<td>[1979Q4]</td>
<td>[1987Q2]</td>
<td>[1989Q2]</td>
<td>[1970Q4]</td>
<td></td>
</tr>
<tr>
<td>Lags</td>
<td>5</td>
<td>4</td>
<td>5†</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Roots</td>
<td>±1,±i</td>
<td>±1,±i</td>
<td>±1</td>
<td>1</td>
<td>1</td>
<td>±1</td>
</tr>
<tr>
<td>Innovative Outlier (IO) model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min $t_{cl}$</td>
<td>-2.5446</td>
<td>-3.7370</td>
<td>-5.7580**</td>
<td>-3.3939</td>
<td>-3.1282</td>
<td>-3.0060</td>
</tr>
<tr>
<td>[1973Q3]</td>
<td>[1979Q2]</td>
<td>[1973Q1]</td>
<td>[1969Q3]</td>
<td>[1973Q3]</td>
<td>[1969Q2]</td>
<td></td>
</tr>
<tr>
<td>Min $t_{c2}$</td>
<td>-2.7205</td>
<td>-4.2547**</td>
<td>-5.8126**</td>
<td>-4.8594**</td>
<td>-4.3275**</td>
<td>-2.3281</td>
</tr>
<tr>
<td>[1973Q4]</td>
<td>[1974Q2]</td>
<td>[1979Q3]</td>
<td>[1979Q4]</td>
<td>[1981Q1]</td>
<td>[1980Q3]</td>
<td></td>
</tr>
<tr>
<td>Max $F_{cl}$</td>
<td>8.7642</td>
<td>8.0794</td>
<td>31.0231**</td>
<td>24.3128**</td>
<td>15.0108**</td>
<td>9.8708</td>
</tr>
<tr>
<td>Lags</td>
<td>5†</td>
<td>4</td>
<td>5†</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Roots</td>
<td>±1,±i</td>
<td>1,±i</td>
<td>5†</td>
<td>1</td>
<td>1</td>
<td>±1</td>
</tr>
</tbody>
</table>

For the first statistic a lag of $k=1$ was considered, since the fifth lag was not significant with a break in 1972Q1 (however, the break-dates and the test results are qualitatively equivalent in both cases); For $F_{cl}$, a lag of $k=4$ was considered, since the fifth lagged variable was not significant with a break in 1970Q4. The asterisks (**) and (***) indicate significance at the 1% and 5% significance levels, respectively. Critical values for $T=120$ (Franses and Vogelsang, 1998): -4.75/-4.78/16.86 (1%), -4.16/-4.19/13.41 (5%) and -3.90/-3.91/11.95 (10%) [AO model], and -4.70/-4.71/16.61 (1%), -4.15/-4.13/12.96 (5%) and -3.86/-3.85/11.49 (10%) [IO model], for the $t_{cl}/t_{c2}/F_{cl}$ statistics, respectively.

The results for the AO model in Table 4.20 reveal that the presence of a long-run unit root can be clearly rejected only in the case of Germany. This suggests that even after allowing for the presence of seasonal mean shifts, the results of the nonseasonal unit root tests keep showing that only Germany satisfies both sustainability criteria. The break-dates generally agree with the findings with the annual data. The only difference is in the case of France, and certainly because the break-date found earlier, 1956, is not included in the shorter period considered with the quarterly data.
For the seasonal unit root tests, the results are only slightly different. The inclusion of the break-point allows the rejection of a unit root at the semi-annual frequency in Germany and at the annual frequency in the United Kingdom. All other previous results of the standard unit root tests hold.

If it is expected that the shifts in the seasonal pattern are gradual, the testing procedure evolves alternatively around the Innovative Outlier (IO) model, and the tests are based directly on the regression

$$\Delta y_t = \pi_1 y_{t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{4,t-1} + \sum_{s=1}^{4} \mu_s D_s + \sum_{s=1}^{4} \delta_s D(UB)_{s} + \sum_{s=1}^{4} \kappa_s D(TB)_{s} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + e_t. \tag{4.12}$$

The test statistics for $t_{11}, t_{12}$ and $F_{34}$ using the IO model are presented in the bottom half of Table 4.20. The methods to choose $k$ and the dates of the breaks are the same as specified above. The results of the tests using this model are very similar to the previous ones. The only difference concerns the unit root tests on the semi-annual frequency in France and at the annual frequency in the United Kingdom, which is basically a consequence of the proximity of the statistics to the critical values.

The break-dates chosen by the tests are also very similar using either the AO or the IO models. For the long run frequency, except for France, all the structural breaks appear to have taken place in the beginning of the seventies. The structural breaks at the seasonal frequencies can be a consequence of methodological adjustments in the national statistical offices, or institutional changes in public accounting, such as changes in tax collection procedures and expenditure allocations, among many other possible reasons. However, in many cases, these seasonal breaks coincide with the breaks in the nonseasonal frequency, suggesting a relation between the seasonal cycle and the business cycle.

After testing the joint condition of sustainability, the next step is to consider separately the first, sufficient condition, which requires a cointegrating relationship between the series of revenues and expenditures. Table 4.21 shows the results of the Engle-Granger and the Johansen cointegration tests applied to the filtered series, as suggested in HEGY.
The model used in the Johansen test includes seasonal dummies, an intercept in the cointegrating vector, and no time trend. The critical values are the same as those used in the standard tests.

Table 4.21: Engle-Granger and Johansen cointegration tests (filtered series)

<table>
<thead>
<tr>
<th>Filters</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$$\delta_t$$</td>
<td>S, A</td>
<td>S, A</td>
<td>-</td>
<td>S, A</td>
<td>A</td>
<td>S, A</td>
</tr>
<tr>
<td>$$\delta_g$$</td>
<td>A</td>
<td>S, A</td>
<td>-</td>
<td>S, A</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Engle-Granger tests

\[
\begin{align*}
\delta_t &= \alpha + \beta \delta_g \\
\delta_g &= \alpha + \beta \delta_t
\end{align*}
\]

<table>
<thead>
<tr>
<th>Filter</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$$\delta_t$$</td>
<td>-1.8183 (7)</td>
<td>-2.4949 (4)</td>
<td>-3.2307* (4)</td>
<td>-1.5328 (0)</td>
<td>-2.4392 (7)</td>
<td>-3.2705* (4)</td>
</tr>
<tr>
<td>$$\delta_g$$</td>
<td>-1.5836 (7)</td>
<td>-1.7221 (4)</td>
<td>-3.5522** (4)</td>
<td>-1.7388 (0)</td>
<td>-2.4890 (7)</td>
<td>-2.9510</td>
</tr>
</tbody>
</table>

Johansen tests

<table>
<thead>
<tr>
<th>Filter</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherl.</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$$\lambda_{MAX}$$</td>
<td>12.4990</td>
<td>8.0194</td>
<td>34.0966**</td>
<td>12.2767</td>
<td>21.4473**</td>
<td>9.6998</td>
</tr>
<tr>
<td>Trace</td>
<td>17.6813</td>
<td>12.5844</td>
<td>48.7357**</td>
<td>16.6260</td>
<td>30.4221**</td>
<td>13.5780</td>
</tr>
<tr>
<td>order</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Notes: S: semi-annual filter, A: annual filter; In parenthesis are the number of lagged dependent variables. The asterisks (**) and (*) indicate rejection of the null hypothesis of no-cointegration at the 5% and 10% significance levels, respectively. Critical values for the EG test, T=159 (MacKinnon, 1991): -3.9675 (1%), -3.3756 (5%) and -3.0720 (10%). Critical levels for the J test (Pesaran, Shin and Smith, 1997): 15.87/20.18 (5%) and 13.81/17.88 (10%) for the $$\lambda_{MAX}/Trace$$ tests and the null hypothesis of $$r=0$$.

The results are similar to the previous ones using annual, central government data. Only in the case of Germany and the Netherlands can the hypothesis of no cointegration be rejected. Furthermore, the tests of the null hypothesis of a unitary cointegrating parameter also confirm the annual results, rejecting the hypothesis in the Netherlands (14.3954***) but not in Germany (0.6510). As before, only Germany satisfies both sustainability conditions, confirming the results of the unit root test on the total surplus in Table 4.19, and also all the previous results using the central government annual data.

In sum, the results of these tests with quarterly data are essentially consistent with the conclusions of the tests with annual data. One important factor to be retained, however, is the necessity to be very cautious when working with quarterly fiscal variables. An apparent relation between the variables may result from inadequately modelled
seasonality. The traditional seasonal adjustment methods often distort the results of the tests, due to the stochastic behaviour of the seasonal pattern of these variables.

Only a few studies have used quarterly government budget variables to examine the question of sustainability of fiscal policies in Europe. MacDonald and Speight (1990) use non-adjusted UK data, but include seasonal dummies in the regressions. They present evidence of a sustainable fiscal policy. Jondeau (1992) found an unsustainable policy in France, using a moving-average (of order 4) method to adjust the data. As observed above, both methods can only account for the deterministic part of the seasonal effect. Caporale (1993a) analysed the fiscal policies of ten EU countries using non-adjusted data and disregarding the seasonality issue, which may explain why he could not find any clear results from the tests (see Table A.I in the appendix).

4.7. European-wide sustainability of fiscal policies

The individual country approach pursued above will be complemented in this section with an analysis at a more aggregate level. The rationale for this section is the argument that with the ongoing process of economic integration, it is the collective situation of the fiscal policies that should be examined (see for example Fitoussi et al., 1993, King, 1995, and Tanzi and Fanizza, 1996). Doubts about the sustainability of a particular country's fiscal policy could affect other countries, even if these follow sustainable policies. It can be argued that an unsustainable fiscal policy in Greece or Portugal, for example, may not adversely affect the major European economies. However, these interdependencies are certainly stronger within the group of countries considered here.

One possible approach to this sustainability issue at a more aggregate level is to consider panel data. Several tests have been proposed recently to examine nonstationary variables with panel data techniques, some of which will be used below to test the existence of a

38 In the US, there are more examples of studies using quarterly or even monthly data. However, these studies generally do not mention the seasonal issues, or use data already seasonally adjusted from the source, and therefore it is usually not possible to evaluate the effects of the seasonal adjustment on the results.
unit root in the series of total surplus. The main advantage of these tests is the significant increase in the number of observations employed. As noted above, a frequent criticism on the standard unit root tests is their low power to distinguish the unit root null against stationary alternatives in small samples. By exploiting cross-sectional information, the panel data unit root tests have higher power.

These tests were first proposed by Quah (1990), but did not attract much interest in the applied work, since they do not allow individual specific effects and the correct consideration of serial correlation in the residuals. The panel unit root tests became very popular only after the papers by Levin and Lin (1992, 1993), henceforth LL, who have proposed a more generally applicable testing procedure. Their approach analyses pooled cross section time series data and tests the hypothesis

\[ H_0: \text{each individual series is I}(1), \text{i.e., } \rho_i = 1 \text{ for } i = 1, 2 \ldots N \]

\[ H_a: \text{all individual series considered as panel are I}(0), \text{i.e., } \rho_i = \rho < 1 \text{ for } i = 1, 2 \ldots N \]

The null hypothesis imposes the cross equation restriction \( \rho = 1 \) on the first order autoregressive coefficients of the equation

\[
(4.13) \quad t_{i,t} = \rho t_{i,t-1} + \bar{\varepsilon}_t,
\]

where the subscript \( i = 1, \ldots, N \) represents the different countries, and \( t_{i,t} \) is the total surplus transformed according to the assumptions made on the stochastic process followed by the different series. The LL test allows for different dynamics across groups and also for common effects. To take account of these, the original data is modified by subtracting from it the individual series' mean and the cross-section mean (to eliminate the influence of aggregate effects).

When there is serial correlation in the idiosyncratic disturbance term \( \varepsilon_{i,t} \), a correction must be made in (4.13), by including lagged differenced terms, as in the ADF tests,

---

39 Tests of cointegration in panel data have also been recently proposed for example by Pedroni (1996 and 1999) and McCoskey and Kao (1998). This line of research was not explored here given the results obtained below.

40 Wu (1996) and MacDonald (1996) apply this methodology to the real exchange rate, to test the PPP hypothesis, Holmes and Wu (1997) to deviations from CIP, Fleissig and Strauss (1997) to real wages, and Culver and Pappel (1997) to inflation.

41 The LL tests assume homogeneity of the autoregressive parameter, which may be a very restrictive assumption as will be seen later.
The test of a unit root in the panel of countries is based on the $t$-statistic for $\rho = 1$ in equations (4.13) or (4.14). Table 4.22 presents the results of these tests for different panels. The first column displays the results for a group of three core members of the European Union: Germany, Netherlands and Belgium. These countries have shown a strong commitment to exchange rate stability and have followed a very similar monetary policy during the last decades. The other columns in the table show the results of the panel data unit root tests when adding alternatively each of the other countries with this core group. Different specifications of the lag length have been employed, although the individual unit root tests in Table 4.2, which provide the benchmark for these results, are unanimous in suggesting no lagged dependent variables.

The authors provide the asymptotic normal distribution of the test statistics. However, the exact finite sample critical values for all these panel-based unit root tests are highly dependent on the number of observations $T$ and the number of individual time series $N$. Therefore, the empirical distribution of the test statistic $t_\rho$ was obtained through Monte Carlo simulations (with 100,000 replications), calibrated for the exact number of countries and observations used in the panel models of Table 4.22.

Table 4.22: Levin-Lin panel data unit root tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-4.3727**</td>
<td>-4.4006*</td>
<td>-4.7491**</td>
<td>-5.3153***</td>
</tr>
<tr>
<td>1</td>
<td>-4.2925**</td>
<td>-4.3654**</td>
<td>-4.5315**</td>
<td>-4.7954**</td>
</tr>
<tr>
<td>2</td>
<td>-4.0910**</td>
<td>-4.4115**</td>
<td>-4.3524**</td>
<td>-4.6216**</td>
</tr>
</tbody>
</table>

The asterisks (**), (*) and (*) indicate rejection of the null hypothesis of a unit root at the 1%, 5% and 10% significance levels, respectively. Critical values (Monte Carlo simulation with 100,000 replications) for three countries: -5.0043/ -4.7052/ -4.625 (1%), -4.2324/ -3.9519/ -3.929 (5%) and -3.8328/ -3.5790/ -3.561 (10%) for 0, 1 and 2 lags, respectively. Critical values for four countries: -5.0995/ -4.8352/ -4.6249 (1%), -4.4108/ -4.1416/ -3.9288 (5%) and -4.0323/ -3.7624/ -3.5613 (10%) for 0, 1 and 2 lags, respectively.

For the core group of countries, the null hypothesis is always rejected whatever the number of lags used in the regression. When another country is added to this core group,
the results continue to show a rejection of the null, but at levels of significance which vary according to the fourth country. When France is included, for example, and model (4.13) is estimated, the statistic is only rejected at the 10% level. This may be explained by the idiosyncratic behaviour of the central government fiscal variables in France, as noted in section 3.2.3.\(^4^2\)

These tests were also applied to different dimensions and permutations of the panel group. The unit root hypothesis is strongly rejected for a panel of all six countries as a whole,\(^4^3\) and for many of the smaller panels, even with only three countries and sometimes even not including Germany in the panel. These results indicate that the non-rejection of the nonstationarity hypothesis is fragile to cross-sectional variation. This raises the question of how much cross-section variation is necessary for the panel tests to reject the null hypothesis of a unit root in the government's total surplus (see Culver and Papell, 1997). Performing these tests on all possible permutations of countries yields the results summarised in Table 4.23 (only the results for zero and one lag are reported). The exact critical values for each group size were computed using Monte Carlo simulations.

Table 4.23: Panel data unit root tests, summary of all possible combinations

<table>
<thead>
<tr>
<th>Countries</th>
<th>Possible combinat.</th>
<th>0 lags</th>
<th>1 lag</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>4 (67%)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>9 (60%)</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>7 (35%)</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>2 (13%)</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The first line reports the results of the individual unit root tests in Table 4.2.

\(^4^2\) These tests were also performed with general government data (for the common period 1953-96), and the null hypothesis of a unit root was always rejected for any of these four-country groups at the usual significance levels.

\(^4^3\) The statistic obtained for the full sample was -5.5716, slightly below the 1% critical value of -5.5125.
The null hypothesis of a unit root is systematically rejected only when five or more countries are grouped together. With less than five countries, there are always some combinations of countries where the null is rejected only at the 10% significance level or higher. The rejection frequencies increase with the sample size. One other finding of these tests is that the rejections happen only when either Germany or the Netherlands are included in the panel, suggesting that these countries act as a type of ‘anchor’ in terms of sustainability.

It has recently been argued that the tests suggested by Levin and Lin are based on the very restrictive alternative hypothesis of homogeneity of the autoregressive parameter \( H_0: \rho_i = \rho < 1 \) for all \( i \). In the particular case analysed here, this implies that if the total surplus is stationary for all countries the coefficient of convergence is the same in every country, a very strong assumption. It may be reasonable to say that, under the null hypothesis, the total surplus does not converge in any country, but it is not reasonable to suggest that, under the alternative hypothesis, the variable will converge in all countries at the same rate. Fleissig and Strauss (1997) suggest that with different coefficients \( \rho_i \) the power of the test is lower. They argue that stationary series should be omitted from the estimation. This would change significantly the results obtained above. Further research is needed to examine the sensitivity of the estimates and critical values to these criticisms, and to extend this test to different specifications of the hypothesis being examined.

More recently, some new tests have been suggested to overcome this problem, allowing heterogeneity of the autoregressive parameter. Im, Pesaran and Shin (1997), for example, henceforth IPS, propose the more realistic alternative hypothesis \( H_a: \rho_i < 1 \) for \( i=1,2,...N \).\(^{44}\) The test is also consistent (when both \( N \) and \( T \) are sufficiently large) with the existence of unit roots on some individual regressions under the alternative hypothesis \( H_a: \rho_i < 1 \) for \( i=1,2,...N_i, \rho_i = 1, i=N_i+1,...,N \).\(^{45}\)

\(^{44}\) Maddala and Wu (1996) propose an alternative, non-parametric test, the ‘Fisher test’, which consists basically of a combination of the significance levels from each individual, independent test, derived by Monte Carlo simulation.

\(^{45}\) Coakley, Kulasi and Smith (1996) and Coakley and Kulasi (1997) apply this test to the relation savings/investment, while Wu (1997), Coakley and Fuertes (1997), and Holmes (1999) employ it to test the hypothesis of a unit root on the real exchange rate.
The basic idea of their test is very simple, and the authors show through stochastic simulations that it has substantially more power than the conventional ADF tests or the LL panel test. Instead of pooling the data, they calculate an average of individual ADF test statistics of the panel

\[(4.15) \quad \bar{t}_{NT} = \frac{1}{N} \sum_{i=1}^{N} t_{iT},\]

where, as before, \(i = 1, \ldots, N\) is the country and \(T\) is the number of observations, and \(t_{iT}\) is the individual \(t\)-statistic for testing \(\rho = 0\) in the ADF regression of the cross-sectionally demeaned series.\(^{46}\) This average \(t\)-statistic is then used to construct the standardised \(t\)-bar statistic\(^{47}\)

\[(4.16) \quad \bar{\Gamma}_i = \frac{\sqrt{N} [\bar{t}_{NT} - E(t_T)]}{\sqrt{\text{Var}(t_T)}} \Rightarrow N(0,1),\]

where \(E(t_T)\) and \(\text{Var}(t_T)\) are the asymptotic values of, respectively, the mean and variance of the average ADF statistic. These values are tabulated in IPS, table 2. Under the null hypothesis, the \(\bar{\Gamma}_i\) statistic has a standard normal distribution, for \(N \to \infty\) and \(T \to \infty\), and for \(N/T \to 0\).

With \(N\) and \(T\) both finite, it is more correct to use simply the statistic \(\bar{t}_{NT}\), which has a non-standard distribution under the null hypothesis that all series are nonstationary. However, the sample sizes tabulated in IPS may not correspond to the actual needs of the test. Furthermore, the IPS’s tables require an equal lag length for all the individual ADF tests, a condition which may be very restrictive. In this case, it is again necessary to compute the exact sample critical values via stochastic simulations, using Monte Carlo methods.

The Levin-Lin panel data test has also been criticised for failing to adequately correct for serial correlation in the model innovations (Papell, 1997). The IPS and Fisher tests are

\(^{46}\) The IPS test implicitly assumes balanced panel data, i.e., the same number of observations for all countries.

\(^{47}\) They also propose an alternative test based on the average of the Lagrange multiplier statistics for each group, called the LM-bar test. However their simulation results show that the \(t\)-bar test performs "marginally better than the LM-bar test" (IPS: p. 2).
based on individual ADF regressions and, since an appropriate ADF regression will correct for serial correlation in the data, these tests take care of serial correlation automatically.48

Table 4.24 reports the results of the IPS test on the same groups of four countries presented for the LL test. The cross-section means were first subtracted from the original data, to remove the time-specific common effect. The tests were again performed using zero to two lags, and also for different lag specifications for each country (0/1).49

Table 4.24: Im, Pesaran and Shin’s panel data unit root tests

<table>
<thead>
<tr>
<th>Lag length</th>
<th>B,G,N and F</th>
<th>B,G,N and I</th>
<th>B,G,N and UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-2.3370**</td>
<td>-2.47011**</td>
<td>-2.6649***</td>
</tr>
<tr>
<td>0/1</td>
<td>-2.1006</td>
<td>-2.3743**</td>
<td>-2.5760**</td>
</tr>
<tr>
<td>1</td>
<td>-2.2090*</td>
<td>-2.1868*</td>
<td>-2.6810***</td>
</tr>
<tr>
<td>2</td>
<td>-2.1905*</td>
<td>-1.9783</td>
<td>-2.4192**</td>
</tr>
</tbody>
</table>

The asterisks (***) , (**) and (*) indicate rejection of the null hypothesis of a unit root at the 1%, 5% and 10% significance levels, respectively. Critical values (Monte Carlo simulation with 100000 replications) for four countries: -2.6618/-2.663/-2.6660/-2.6279 (1%), -2.3212/-2.320/-2.3191/-2.2931 (5%) and -2.1444/-2.145/-2.1463/-2.1136 (10%) for 0, 0/1, 1 and 2 lags, respectively.

Using Monte Carlo simulations, IPS claim that their test has better finite sample properties than the LL test. However, in the particular case under study, the results of the IPS test basically confirm the results previously obtained. The null hypothesis of a unit root is rejected at different levels of significance, according to the fourth country added to the core group of three countries. With the heterogeneous lag specification suggested by the individual tests, the null is not rejected even at the 10% significance level when France is included.

The panel data methodologies to deal with nonstationary variables are still in a very early stage of development, and therefore the results of all these tests must be more carefully

48 In strict terms, the IPS and Fisher tests are not exactly panel unit root tests, unlike the LL test which is based on panel estimation. They just gather information from several individual ADF tests and combine, respectively, their t-statistics and significance levels.
49 One lag for Germany and the United Kingdom, and zero lags for all others, following the indications of the individual tests.
interpreted. The existence of contemporaneous correlation between the countries, for example, may cause size distortions in the tests discussed above, mainly the IPS and Fisher tests, which are averages of independent $t$-statistics and independent $p$-values (see for example Wu, 1997). The LL test is basically a SURE procedure and therefore, as noted by Maddala (1997), it may allow for contemporaneous correlation across countries, depending on the relative dimension of the contemporaneous correlation matrix relatively to the number of observations.$^{50}$

One other potential problem with these panel unit root tests is their sensitivity to the presence of outliers in the group being analysed (see, for example, Taylor and Sarno, 1998). The IPS test statistic, for example, being an average of several individual tests, may be influenced by an extremely high or low $t$-statistic of one of the individual regressions. This may be particularly serious for very small panels, as the ones used in this section. Therefore, rejection of the null hypothesis that all series are nonstationary does not necessarily indicate that all series are stationary. These tests are also sensitive to the number of individual series considered. The power of the panel unit root tests is highly dependent on the extent of cross-sectional variation, and therefore on the number of countries included in the panel, as shown in Table 4.23.

For all these reasons, these tests do not allow inference concerning a particular country, they must be regarded in more general, conceptual terms. In this sense, the results of the panel data unit root tests cannot be directly compared with the results of the standard unit root tests in individual countries. The hypothesis being tested is different. As a general hypothesis, these tests indicate that the total surplus is a stationary variable when a sufficient number of countries is considered, and therefore, in general the fiscal policy in this group of countries is sustainable.

An alternative, more straightforward method of analysing the question of a European-wide sustainability of fiscal policies, is simply to aggregate all revenues and expenditures of this group of countries and repeat the standard procedure. The hypothesis being studied is the following. Suppose that a central fiscal authority is set in EMU, as happened with the monetary policy. The European-wide budget would basically be a

$^{50}$ However, O'Connel (1998) reports that also the LL test may suffer from severe size distortions,
consolidation of the national budgets.\textsuperscript{51} Would this European-wide fiscal policy be sustainable?

Table 4.25 presents the results of the unit root and cointegration tests on the consolidated variables of the six EU central governments. To obtain the aggregate series, as ratios of GDP, the value of all national variables was first transformed from domestic currency into dollars.

Table 4.25: Unit root and cointegration tests on the aggregate fiscal variables

<table>
<thead>
<tr>
<th>ADF\textsubscript{SBC} unit root test</th>
<th>$t^A$</th>
<th>$t^g$</th>
<th>$t^s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1.0059</td>
<td>-1.1155</td>
<td>-2.0255</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
</tbody>
</table>

Johansen cointegration test (order=1, model with a restricted intercept and not trend)

<table>
<thead>
<tr>
<th>$\lambda_{\text{MAX}}$</th>
<th>$\text{Trace}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>11.8370</td>
</tr>
<tr>
<td>$H_1$</td>
<td>1.3603</td>
</tr>
<tr>
<td></td>
<td>1.3603</td>
</tr>
</tbody>
</table>

In the ADF\textsubscript{SBC} the lag length is chosen according to the 'Schwarz Bayesian Criterion'. In parenthesis is the number of lagged differenced terms used in the regression. Critical values of the DF\textsubscript{SBC} test (MacKinnon, 1991): -2.60 (10%), -2.93 (5%) and -3.58 (1%). Critical levels of the Johansen test (Pesaran, Shin and Smith, 1997): 15.87/20.18 (5%) and 13.81/17.88 (10%) for the $\lambda_{\text{MAX}}$ Trace tests and the null hypothesis of $r=0$, and 9.16 (5%) and 7.53 (10%), for both tests of $r\leq 1$.

The generalised conclusion from these tests is that the aggregate fiscal policy of the six EU members is not sustainable, suggesting the need for policy coordination. This result contradicts the conclusions from the panel unit root tests, not allowing final, definite conclusions. This may be due to the fact that simple aggregation involves the use of exchange rates which may distort the results. It may also be a consequence of the inevitable trade-off between the low power of the standard nonstationarity tests, and the size distortions of the panel unit root tests.

\textsuperscript{51} This is equivalent to the consolidation of the regional and local governments' accounts, to obtain the general government's budget. Cooper (1990: p. 280), for example, proposes that all EU public debt should be consolidated and become a liability of the EU as a whole. Jordan (1997: p. 5) argues that "the consolidated fiscal positions of all the member countries will affect confidence in the soundness of the currency".
4.8. Summary of the main results

This chapter has tested the hypothesis of sustainability of fiscal policies in six EU member-countries. The analysis has proceeded along a sequential strategy which allows not only the verification of sustainability in the strict sense, but also the complementary condition of a bounded debt-GDP ratio. The overwhelming evidence suggests that the central government's fiscal policy in Germany and the Netherlands can be considered sustainable, but the latter country may still face the problem of a continuous increase of the debt ratio. Limits on the capacity or willingness of the private sector to absorb expanding quantities of government bonds may lead to an increase in the interest rate demanded by the markets. This could rapidly turn a sustainable into an unsustainable fiscal position, given the dynamics of the IBC. The results for the UK are not unanimous, but overall are more indicative of an unsustainable policy. For all the other countries in the sample, the results of the tests strongly suggest that policy is not sustainable, pointing the need for a change in the current direction of their fiscal policies.

These results were shown to be robust to structural breaks in the data. For most countries the tests reveal a mean shift occurring in the beginning of the seventies, usually justified by the expansionary fiscal policies following the first oil shock. However, other factors may have also contributed to this structural change, such as the increased autonomy of national economic policies after the end of the Bretton Woods system of fixed exchange rates.

Previous empirical work on the sustainability hypothesis have suggested that the results may also be sensitive to, among other factors, the level of government accounts considered and the frequency of the observations. Both questions have been examined thoroughly in this chapter, and the overall inference is that the initial main general conclusions are robust to these factors. Only a few points are worth reporting. First, the increasing influence of the social security imbalances on the reported deficits of the general government. Although the deficits of the central authority have been on average above those of the general government throughout the sample period considered here, this situation seems to have changed by the end of the sample, with the deficits of the social security funds dominating those of the central, regional and local authorities. This
problem was particularly notorious in Germany, possibly the main reason why the second condition of a stable debt ratio is not satisfied with general government data in this country, while it is in the Netherlands. Second, the problems of working with fiscal data at frequencies higher than the annual. The seasonal adjustment methods are either ineffective or potentially misleading with these variables. However, after testing for, and incorporating in the model, the possibility of nonstationary seasonality at the various frequencies, the analysis delivered basically the same conclusions previously obtained with the homologous annual data set.

As a complement to the sustainability analysis, the chapter has also considered the issue of causality between the budget variables. In order to obey the IBC, a fiscal disequilibrium must be offset either by increasing revenues, falling expenditures or a combination of both. This question has received large attention recently, in both theoretical and empirical terms, without a definite conclusion having been reached. The tests performed with central government data indicate that in Germany the fiscal imbalances have been predominantly offset by cuts in spending, while in the Netherlands the burden of adjustment falls generally on the revenue side. This may help explain why the share of the central government in the economy has substantially increased in the Netherlands during the last decades in comparison with Germany, as a result of this spend-and-tax policy.

The causality analysis has a particular importance for the group of countries with unsustainable fiscal policies, as an indication of the more efficient strategy to change the fiscal stance. However, within this group only for Italy could a significant result be found, revealing that revenues Granger-cause expenditures, that is, an increase in taxes to curtail the deficit, for example, may be ineffective since it is followed by higher spending. This suggests that a more correct fiscal consolidation policy in this country should put an emphasis on spending cuts.

When general government data is used, the non-causality tests indicate a predominant effect running from expenditures to revenues in France, Germany and the Netherlands. This may be explained by the lower elasticity of the social security expenditures, and justifies the increasing concerns with the fiscal burden arising from the expected demographic evolution in most industrial countries.
Finally, this chapter has also explored the issue of sustainability at the aggregate level, either by using panel data techniques or by consolidating the fiscal variables into a 'supranational budget'. The panel data unit root tests have revealed that if a sufficient number of countries is included, both sustainability conditions are satisfied, in general terms. Opposite conclusions were obtained when repeating the sequential testing strategies on the consolidated variables. A hypothetical supranational fiscal authority could be faced with an unsustainable fiscal policy. The conclusions of this section must be carefully regarded, given the small number of countries considered and, therefore, the sensitivity of the results to extreme values in the sample.

After having detected problems of unsustainable fiscal policies in most of the EU countries analysed, the immediate question to consider is the possible economic consequences of these policies. The next two chapters will examine this question empirically, focusing particularly on the effects of these fiscal policies on monetization, inflation and the interest rate.
5. MONETIZATION OF THE DEBT AND DEFICITS

5.1. Introduction

The analysis of sustainability of fiscal policies is largely motivated by the concern that an unsustainable fiscal situation inevitably implies the monetization of the debt and deficits, ultimately resulting in higher inflation (Committee for the Study of EMU, 1989, Lamfalussy, 1989, Commission of the EC, 1990, Buiter, 1995b, and others). However, the empirical studies of sustainability usually abstract from the effects of money creation on the government’s intertemporal budget constraint (IBC), considering them to be negligible. Bohn (1991a: p. 336), for example, considers that “it would only be distracting to treat the inflation-tax as a special case”.

Some authors acknowledge the importance of this source of revenues in certain countries. Heinemann (1993), for example, attributes the responsibility for the contradictory results in his sustainability study of the Italian fiscal policy (see Table A.1 in appendix) to the neglected importance of seigniorage as a means of budget financing. Baglioni and Cherubini (1993) suggest that the present Italian unsustainable fiscal situation is due to the loss of seigniorage revenues since the early eighties. However, neither paper attempts to verify empirically these suspicions. Uctum and Wickens (1996) analyse the effect of inflation on sustainability, but only within a finite horizon framework, using projections of the variables for the future. They perform simulations with the primary deficit unadjusted for seigniorage, and found a non-negligible effect in some European countries.

Trehan and Walsh (1988) present the only known study of long-run fiscal sustainability to address this specific question. They isolate the seigniorage revenues from all other sources of government income, testing the hypothesis of non-cointegration between the series of total government expenditures, tax receipts and seigniorage in the US. However,
the results obtained are extremely ambiguous. They find that both series of total deficit, with and without seigniorage revenues, are stationary, which is not consistent with their previous finding of a unit root in the seigniorage series.

This chapter will estimate the relative importance of the seigniorage revenues in the European countries, and investigate whether they have played a significant role as an alternative financing source. It will further examine whether budgetary disequilibriums necessarily imply monetization and inflation. The chapter is organised as follows. The next section reviews the approaches usually followed to analyse the connection between the government's budget dynamics and the levels of monetization and inflation. Section 5.3 presents the theoretical definition of seigniorage, and the main methodological approaches for its empirical estimation. Both measures of seigniorage computed here are employed in section 5.4 to investigate whether the EU governments have been obeying the 'non-monetization budget constraint'. Section 5.5 provides some empirical evidence on the relationship between the budget deficit and the time series of seigniorage and inflation. Section 5.6 sets out the underlying model of public finance describing the relation between seigniorage and other sources of government revenues, derives its theoretical implications, and tests them empirically. Finally, section 5.7 summarises and offers some concluding remarks on the results of this chapter.

5.2. Theoretical background

A variety of different approaches have been suggested in the literature to examine the relationship between fiscal and monetary policies or, more concretely, between the public finances and the levels of seigniorage and inflation. One approach uses the analytical framework proposed by Sargent and Wallace (1981), where seigniorage assumes the central role for deficit finance. It views this relation basically as a game between the fiscal and the monetary authorities. Under the assumption that the fiscal authority is the first to make a move (is the 'Stackelberg leader'), the monetary authority is left with a difficult choice in order to balance the intertemporal budget (the 'unpleasant monetarist arithmetic'). It either loosens its policy in the short-run to avoid high inflation in the long-
run, or tightens its policy today at the expense of an inevitable future increase in inflation. The tighter the monetary policy today, the worse are the fiscal problems. A tight monetary policy increases the real interest rate (and consequently the debt burden), slows the economy (worsening the deficit through the automatic stabilisers), and reduces the seigniorage revenues. Sooner or later, the fiscal imbalance must be financed through seigniorage. In sum, a fiscal dominant regime implies a positive intertemporal correlation between deficits and money growth.

However, if the monetary authority is allowed to move first, it could ensure price stability and force the fiscal authority (the ‘Stackelberg follower’) to obey the IBC through fiscal consolidation. This idea is behind the institutional changes designed to provide the central banks with more independence, a sufficient condition to avoid monetization, according for example to Buter et al. (1993) and King (1995), not sufficient according to Von Hagen (1998) or Woodford (1998).

All the above arguments of either a fiscal or a monetary dominant regime assume that the IBC necessarily holds, and therefore either the monetary or the fiscal authority have to adjust to whoever makes the first move in the ‘game’. However, as empirically shown in the previous chapter, the IBC is not satisfied in most EU countries. It is possible that the monetary and fiscal decisions are taken independently, potentially resulting in unsustainable policies.

The conjecture first advanced by Sargent and Wallace (1981) relies on seigniorage revenues to balance the budget. Many economists claim that this situation is very unlikely in a country traditionally presenting a low rate of inflation (see, for example, King, 1995). One other approach to the link between fiscal and monetary policy argues that the main channel is the effect of inflation on the real value of the stock of debt and on the real interest rate (see, for example, Woodford, 1998, or Dombusch, 1998). While Sargent and Wallace (1981) saw the first approach as an economic policy game between fiscal and monetary authorities, Keynes (1971), for example saw this second as a ‘social game’ between rentiers and workers. Society does not tolerate ever-increasing taxes, and therefore other ways, such as inflation, must be considered to reduce the accumulated stock of debt.
This link is especially important in the exceptional context of hyperinflation, when debt can be almost completely devalued. At present, this option is certainly less likely in most developed economies, given that the stock of debt is increasingly more composed of short-term instruments (Missale and Blanchard, 1994, and De Grauwe, 1996). The investors anticipate the temptation to devalue the debt and include these expectations in the nominal interest rate demanded. In fact, although inflation may not be used to reduce the debt value, if the monetary authority is completely unresponsive to its size (no monetization), high debt can still contribute to higher inflation, through expectations and also perhaps through a wealth effect on private spending decisions; although the Ricardian Equivalence Theorem would suggest that the latter channel is negligible.

These two main approaches highlight the differences between expected and unexpected inflation, and their effects on the budget. While expected inflation increases the seigniorage revenues, unexpected inflation reduces the real debt burden. This difference can be easily demonstrated by looking at the government’s budget constraint, and separating the elements which are sensitive to the inflation rate. Starting from the government’s budget identity introduced in chapter 2, equation (2.8), with the variables presented as ratios to GDP,

\[
(5.1) \quad d_t = \frac{(1 + i_t)}{(1 + \psi_t)(1 + \pi_t)} d_{t-1} + g_t - t_t - \left[ m_t - \frac{m_{t-1}}{(1 + \psi_t)(1 + \pi_t)} \right]
\]

where, as before, \( d_t \) represents the real stock of debt at the end of the period, \( g_t \) and \( t_t \) indicate, respectively, the real primary expenditures and revenues, \( i_t \) is the nominal interest rate, \( \psi_t \) the growth rate of the economy, and \( \pi_t \) the rate of inflation. With price stability, this equation would reduce to

\[
(5.2) \quad d_t = \frac{(1 + i_t)}{(1 + \psi_t)} d_{t-1} + g_t - t_t - \left[ m_t - \frac{m_{t-1}}{(1 + \psi_t)} \right].
\]

To identify the price or, inflation effect (IE) on the budget identity, equation (5.1) is subtracted from equation (5.2), yielding

\[
\text{IE} = \frac{(1 + i_t)}{(1 + \psi_t)} d_{t-1} - \left[ m_t - \frac{m_{t-1}}{(1 + \psi_t)} \right].
\]
This expression displays the effect of inflation on the two main nominal liabilities of the public sector, the public debt and the monetary base. It represents the public revenues resultant from the so-called ‘inflation tax’. Since the devaluation of public debt can be completely offset by a perfectly anticipated rate of inflation, the second term on the right-hand side of (5.3) is sometimes referred to as the ‘expected inflation tax’ (see, for example, Buiter, 1985).

The first term on the right-hand side of (5.3) gives the devaluation of the stock of interest bearing public debt, and acts, in practice, as an amortisation payment (Buchanan, 1996a). By subtracting this term from the total deficit, the so called ‘Operational Deficit’ is obtained, defined as the primary deficit plus the revenues from money base creation, increased by the real component of interest payments.

To take the inflation effect out of the debt service, and of the primary deficit, the budget identity (5.1) may be represented as

\[
IE_t = \frac{(1 + i_t) \pi_t}{(1 + \psi_t)(1 + \pi_t)} d_{t-1} + \frac{\pi_t}{(1 + \psi_t)(1 + \pi_t)} m_{t-1}.
\]

The right-hand side of equation (5.4), excluding its first term, is the ‘inflation corrected primary deficit’, including seigniorage revenues. The fourth term on the right-hand side gives the ‘inflation effect’, the change in the real value of the debt due to changes in the price level. This inflation effect is zero only with a zero inflation rate. By subtracting this term from the public deficit, the inflation tax is being included in the public revenues (see, for example, Eisner and Piepper, 1984).

---

2 Other possible effects of inflation on the primary balance are not being considered in this analytical framework, namely the effect of inflation on the government’s expenditures and revenues, more relevant for high inflation countries. Therefore, to isolate the inflation effect, it will be considered for now that the budget variables, ceteris paribus, grow proportionally to inflation.

3 See, for example, Tanzi et al. (1988) and Blejer and Cheasty (1991).
If the real interest rate is invariant with inflation and if the Fisher hypothesis holds, i.e., if inflation is correctly anticipated by the economic agents (or if debt is completely indexed to the price level), the nominal interest rate would have incorporated the inflation tax rate, and there would be no net gain for the government from the devaluation of its debt (see, for example, Buiter, 1983). This can be more clearly analysed by further decomposing the public sector's budget identity and including the Fisher identity $(1+i_t) = (1+r_t)(1+\pi_t^e)$,

$$d_t = \frac{(1+r_t)}{(1+\psi_t)} d_{t-1} + \left[ \frac{(1+r_t)\pi_t^e}{(1+\psi_t)} - \frac{(1+r_t)(1+\pi_t^e)\pi_t}{(1+\psi_t)(1+\pi_t)} \right] d_{t-1} + g_t - t_t - \Delta m_t - \frac{\psi_t}{(1+\psi_t)} m_{t-1} - \frac{\pi_t}{(1+\psi_t)(1+\pi_t)} m_{t-1}.$$

This equation reveals that if inflation is perfectly anticipated by the investors, and included in the nominal interest rate, the term in square brackets is zero. In this situation, the real stock of interest-bearing debt will only be altered if inflation affects the real interest rate, the primary deficit, or the real revenues from base money creation. The government can only obtain net revenues from the devaluation of its real stock of debt through a surprise inflation ($\pi_t > \pi_t^e$) or if the nominal interest rate is artificially constrained. However, expected inflation still influences the budget, by increasing the amount of seigniorage revenues.

The term $(\psi_t m_{t-1}) / (1+\psi_t)$ indicates the amount of base money creation induced by the growth rate of the economy. If the government just wanted to maintain a constant ratio of base money to GDP, i.e., a constant velocity of circulation of base money, it could nevertheless increase the nominal value of base money in the proportion given by this term, in order to accompany the growth of the economy.

The last term in equation (5.5) displays the effect of inflation on the monetary base, and gives the proportional increase in the budget surplus due to a devaluation of the real monetary base. However, higher inflation usually leads to higher nominal interest rates and possibly to a lower demand for real money balances, raising the velocity of

---

4 Evidently, if the Fisher hypothesis does not hold, and the real interest rate is reduced by higher anticipated inflation, then the real stock of debt devalues even due to higher anticipated inflation.
circulation of the monetary base. This implies that the term $\Delta m$, may be negative, compensating part of the inflation effect on money balances (see Figure A.8 of the appendix, discussed below).

The correction of the public deficit for the effects of inflation may have some impact on the empirical analysis of sustainability. By clearly identifying the different terms in equation (5.5), it is possible to investigate the importance of base money creation and of inflation on the sustainability of fiscal policies. This hypothesis will be investigated in the next sections.

However, other channels can be identified. One is the wealth effect, a demand-side effect which can be found for example in Patinkin (1965). The rise in the real value of the stock of bonds increases the perceived private wealth and therefore spending, leading to inflation. There may be also a supply-side effect, through the cost of factors. A deficit caused by an increase in certain public expenditures may augment the demand for scarce resources, increasing their cost (see, for example, Wray, 1997). These two effects may be regarded as the 'non-monetization effects' of deficits.

The wealth effect was particularly denied by Barro's Ricardian Equivalence. Instead, Barro (1979) focus on the inverse positive causal relation. High inflation expectations ($\pi^e > \pi$) imply higher real interest rates and therefore an heavier debt burden, an important component of the deficit in highly indebted countries.

Finally, a different view of this connection between monetary policy (inflation) and fiscal policy (debt and deficits) is proposed for example by Persson et al. (1998). According to these authors, the fiscal gains from higher inflation arise mainly from the nominal features of the tax/transfer system, rather than from the more traditional sources considered above of seigniorage and devaluation of real debt. In general, inflation tends to increase the real, effective tax rates while eroding the real value of the money transfers, due for example to an imperfect indexation of the tax and transfer systems.\(^5\) Analysing the particular case of Sweden in 1994, they conclude that the fiscal gains from inflation may be quite substantial, although the welfare costs may be even higher. The problem

\(^5\) This effect contradicts the more conventional Tanzi (or Olivera-Tanzi) effect of the nominal characteristics of the tax system, according to which higher inflation reduces the real government revenues
may be that the political costs are lower than the social costs. The absolute fiscal gains from inflation should therefore be compared with its welfare costs. This discussion is postponed to section 5.6, with the public finance models of optimal revenue collection, which represent a first approach to analyse these costs.

Since the late seventies, several studies, focused almost exclusively on the US case, have attempted to verify empirically the existence of the above effects. The results have been far from consensual. On the one hand, Allen and Smith (1983), McMillin (1986) and Hondroyiannis and Papapetrou (1997), among others, report some evidence that deficits have an effect on money growth and inflation. On the other hand, Barro (1987), Joines (1990), King (1995) and Abizadeh and Yousefi (1998), for example, argue that such relation is not found in the data. Dwyer (1982) and Joines (1985) present evidence that higher inflation contributes to larger government deficits. Smithin (1994) suggests that growing deficits are historically connected with lower inflation rates. King (1998) examines this question with a very simple pooled regression of inflation on debt, for several OECD countries, and also finds evidence of a negative relation.

5.3. Measurement of the seigniorage revenues

In a broad sense, seigniorage is the income accruing to the public sector as a result of its sovereign power to issue base money. In other words, the public sector obtains seigniorage revenues when it sells very low or interest-free financial assets to the private sector. These revenues are obtained indirectly from the central bank, through the government’s access to the monetary base to finance its deficits. This general definition of seigniorage is relatively consensual. However, the question of what is the most correct measure of the seigniorage revenues obtained by the government has not yet been settled. Two main measures have been developed in the literature and will be successively presented and estimated in this section.
The measure of seigniorage most commonly considered is the so-called 'monetary' or 'cash-flow' definition of seigniorage, given simply by the change in the monetary base during a certain period. It measures the flow of government expenditures financed by the emission of extra amounts of currency and the use of additional amounts of reserves of the banking sector. This process of money creation can be decomposed in three components, the last three terms in equation (5.5)

\[
(5.6) \quad \sigma^F = \frac{(M^N_t - M^N_{t-1})}{GDP_t} = \Delta m_t + \frac{\psi_t}{(1 + \psi_t)} m_{t-1} + \frac{\pi_t}{(1 + \psi_t)(1 + \pi_t)} m_{t-1},
\]

where \( \sigma^F \) indicates the amount of seigniorage revenues as a share of GDP according to the cash-flow definition. This equation shows clearly that the total amount of money creation depends directly on the public's real money demand, the growth rate of the economy, and the inflation rate. The sum of the two last terms in (5.6) approaches unity when inflation approaches infinity. To observe the relative importance of each of these components see Figure A.8 in the appendix. As noted above, the first component presents negative values during certain periods in all countries, and is very volatile. The third component presents very large values during the mid-seventies, impelled by the high inflation rates. However, all components, in all countries, converge to a very low value by the end of the sample, revealing the declining importance of this source of revenues for the government.\(^7\)

Aggregating all three components, Figure A.9 in the appendix displays the total amount of seigniorage revenues appropriated by the government, relative to GDP, according to the cash-flow measure (the darker line). Italy appears as an exceptional case, presenting very high values of seigniorage throughout the seventies, a period of significant monetization of the Italian government deficits (Fratianni and Spinelli, 1999). Although this suggests that countries which rely more significantly on seigniorage also present high public debts, the inverse conclusion cannot be inferred. In other highly indebted countries such as Belgium and the Netherlands, the average values of seigniorage are almost

\(^7\) Grilli (1989: p. 73) argues that the importance of seigniorage may increase in the future. The unsustainable fiscal policies in many countries require important fiscal adjustments in which seigniorage may be considered: "While it is true that inflation does not need to be part of a fiscal reform, the existence of domestic political constraints on the use of alternative sources or revenues may make seigniorage indispensable".
negligible. This variable is highly volatile in all countries, and presents negative values in certain periods. This is due to the high sensitivity of this measure to slight changes in the currency and reserves ratios, reflecting the demand for real money balances and the reserves policy of the banking sector.

An alternative measure, also shown in Figure A.9 (grey line), is known as the 'opportunity cost' definition of seigniorage.\(^8\) It is calculated as the product of the total amount of monetary base and the average interest rate on the public debt. While the previous definition is based on a flow concept, and is therefore more volatile, this one is based on a stock definition. The underlying idea is that the monetary base is a liability of the public sector, on which no (or low) interest is paid, and therefore seigniorage is the amount of interest payments saved by the government. In practice, what is transferred to the government are the central bank's profits, which are not equivalent, in a particular period, to the change in the monetary base. The government's budget should then be represented by

\[
D_t = (1 + i_t)D_{t-1} + G_t - T_t - \Pi_t
\]

where \(\Pi_t\) symbolises the profits of the central bank. Consider an extremely simplified balance sheet of the central bank

\[
A_t + D_t^{CB} = M_t
\]

where \(D_t^{CB}\) is the government's debt in the central bank, \(A_t\) represents all the other central bank's assets, and \(M_t\), the monetary base, are the total liabilities. Two further assumptions have to be made (see, for example, Repullo, 1991). First, the interest rate earned by the central bank on all its assets is the same, \(i_t\).\(^9\) Second, it is assumed that no interest is paid on commercial bank's reserves, which may be suspect in some cases, influencing the results of tests using this measure. In this situation,

\[
\Pi_t = i_t(A_t + D_t^{CB}) = i_t M_t = \sigma^{OC}, \text{ and,}
\]

\(^8\) This was the measure used for example in the most quoted report 'One market, one money', presented by the Commission of the EC in 1990.

\(^9\) If the interest rate on \(A_t\) is below \(i_t\), it would be profitable for the commercial banks to get credit from the central bank to buy government debt. On the other hand, more realistically, if the interest rate on \(D_t^{CB}\) is below \(i_t\), the lower interest payments on the public debt would be exactly compensated by the reduction in the profits of the central bank, not affecting the budget constraint (Baltensperger and Jordan, 1997).
The main practical problem with this concept of seigniorage is the choice of the appropriate interest rate, which introduces some arbitrariness in the empirical studies. The long-term interest rate on government bonds is usually the preferred 'proxy', and it was also used here to estimate the series presented in Figure A.9. As can be observed, this measure is based on more stable variables.

Table 5.1 below shows that, except for Belgium, the average value of seigniorage obtained by the government is very similar, whichever measure is used. In fact, in the long-run these two measures of seigniorage are intertemporally equivalent (see, for example, Buiter, 1995a). Formally, the present-value of the future flows of seigniorage revenues relative to GDP can be represented, for each measure of seigniorage, as

\[
PV(\sigma^{CF}) = \sum_{j=0}^{\infty} \left[ \prod_{k=1}^{j} \frac{(1 + i_{t+k-1})}{(1 + \pi_{t+k-1})(1 + \psi_{t+k-1})} \right]^{-1} \sigma_{t+j}^{CF}, \quad \text{and}
\]

\[
PV(\sigma^{OC}) = \sum_{j=0}^{\infty} \left[ \prod_{k=1}^{j} \frac{(1 + i_{t+k-1})}{(1 + \pi_{t+k-1})(1 + \psi_{t+k-1})} \right]^{-1} \sigma_{t+j}^{OC}.
\]

Denoting the discount factor by \( \delta_{t,j} \) to simplify the notation, equations (5.11) and (5.12) can be algebraically modified to appear as

\[
\sum_{j=0}^{\infty} \delta_{t,j} \sigma_{t+j}^{CF} = \sum_{j=0}^{\infty} \delta_{t,j} m_{t+j} - \sum_{j=0}^{\infty} \delta_{t,j} \frac{m_{t+j-1}}{(1 + \pi_{t+j})(1 + \psi_{t+j})}, \quad \text{and}
\]

\[
\sum_{j=0}^{\infty} \delta_{t,j} \sigma_{t+j}^{OC} = \sum_{j=0}^{\infty} \delta_{t,j} \frac{i_{t+j} m_{t+j-1}}{(1 + \pi_{t+j})(1 + \psi_{t+j})}.
\]

Subtracting (5.14) from (5.13) yields

\[
PV(\sigma^{CF}) - PV(\sigma^{OC}) = \sum_{j=0}^{\infty} \delta_{t,j} m_{t+j} - \sum_{j=0}^{\infty} \delta_{t,j} m_{t+j-1} = -m_{t-1}
\]

\[ ^{10} \text{Other interest rates could have been used as proxies of the average interest rate on the government debt. Buiter (1995a), for example, uses TB rates and call money rates (lines 60b and 60c of the IFS).} \]
The difference between the present-value of both measures of seigniorage revenues is given by the beginning of the period stock of base money relative to GDP. Therefore, it is expected that the seigniorage revenues measured by the 'opportunity cost' definition provide very slightly higher revenues for the government, negligible in a very long period. In fact, Table 5.1 demonstrates this for the countries being studied here. With the exception of Italy, in all other countries the average value of seigniorage measured by the 'opportunity cost' definition is higher than when measured by the alternative definition, even though only a limited time span is considered.

Table 5.1: Average value of seigniorage during the sample period (% GDP)

<table>
<thead>
<tr>
<th>Seigniorage measure</th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash-flow</td>
<td>0.580</td>
<td>0.741</td>
<td>0.745</td>
<td>1.642</td>
<td>0.583</td>
<td>0.443</td>
</tr>
<tr>
<td>Opportunity cost</td>
<td>1.036</td>
<td>0.753</td>
<td>0.762</td>
<td>1.535</td>
<td>0.618</td>
<td>0.609</td>
</tr>
<tr>
<td>Correlation coeff.</td>
<td>0.3869</td>
<td>0.5249</td>
<td>0.1499</td>
<td>0.3886</td>
<td>0.0632</td>
<td>0.4271</td>
</tr>
</tbody>
</table>

In spite of the long-run equivalence, the behaviour of the two series in the short-run can be substantially different for a particular time period, as can be observed in Figure A.9 of the appendix. The cash-flow definition can even show negative values when the monetary base is very volatile. The correlation coefficient between the two measures is usually inferior to 0.5, as shown in the bottom line of Table 5.1, which demonstrates that both concepts originate diverging movements in the evolution of the series. However, as argued by Gros (1993) for example, there is no clear theoretical justification to choose a particular concept. Therefore, both measures will be alternatively used in the next sections to evaluate the importance of these revenues on the sustainability analysis.

---

11 Identical values were obtained by Dupuy (1993), for example. Honohan (1996) even found a negative value of the correlation coefficient between both seigniorage measures in Greece.
5.4. The non-monetization intertemporal budget constraint

Although still significant in some countries, the prevalent worldwide trend in recent years has been for a decline in the relative importance of seigniorage. This process is expected to continue in the future, given for example the current tendency to strengthen central bank independence with the primary objective of ensuring price stability. In a monetary union, for instance, national authorities lose the ability to choose independently the preferred inflation rate, and the residual seigniorage receipts are collected by the central authorities and redistributed by the member-countries according to some rule.  

The main purpose of this section is to verify whether past seigniorage revenues have played a significant role in defining the (un)sustainability of fiscal policies. In particular, the hypothesis to test is: if governments had not benefited from seigniorage proceeds during the post-war period, and assuming that the evolution of the other revenues and of all expenditures were not affected by the unavailability of this source of income, will the sustainability tests present different results? In other words, do EU governments obey the non-monetization constraint or, as De Grauwe (1992) designates it, the 'harder budget constraint'?

Table 5.2 summarises the results of the sustainability tests on central government data. The unit root analysis (not shown) on the underlying revenue and expenditure variables, net of seigniorage revenues, did not significantly alter the results previously obtained in chapter 4, allowing the application of the cointegration tests. The table presents the results of the ADF\(_{sbc}\) unit root tests on the series of the net-of-seigniorage surplus (\(ts_{r} - \sigma_{r}\)), and the Johansen tests of cointegration between total expenditures (\(tg_{t}\)) and net-of-seigniorage revenues (\(tt_{r} - \sigma_{r}\)). Other econometric techniques strengthened the results but are not shown to avoid ‘overloading’ the table. Where the two series were found to be cointegrated, a Wald test on the cointegrating parameter examines whether the supplementary sustainability condition of a bounded debt ratio is satisfied. The top panel

---

\(^{12}\) In the particular case of EMU, these revenues are redistributed according to the participation of each country in the capital of the European Central Bank. In turn, these depend on the relative share of each country's population and GDP ('Protocol on the Statute of the European System of Central Banks and of the European Central Bank', Article 35, no. 5).
of the table presents the results when the cash-flow measure of seigniorage is used, while the bottom panel refers to the opportunity cost definition.

### Table 5.2: Sustainability tests on the non-monetization constraint

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unit root test on the net-of-seigniorage ($\sigma^{CF}$) surplus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF$_{SRC}$</td>
<td>-1.5738</td>
<td>-2.3808</td>
<td>-5.8494</td>
<td>-1.0995</td>
<td>-3.0862</td>
<td>-2.8606</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(2)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td><strong>Cointegration test between total expenditures and net-of-seigniorage ($\sigma^{CF}$) revenues</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\lambda_{MAX}) test</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(H_0: r = 0)</td>
<td>5.7472</td>
<td>16.8964</td>
<td>27.3045</td>
<td>15.0014</td>
<td>17.6168</td>
<td>10.8392</td>
</tr>
<tr>
<td>(r \leq 1)</td>
<td>4.4309</td>
<td>3.8016</td>
<td>7.3271</td>
<td>1.9034</td>
<td>2.6112</td>
<td>2.3006</td>
</tr>
<tr>
<td><strong>Trace test</strong></td>
<td>10.1781</td>
<td>20.6980</td>
<td>34.6316</td>
<td>16.9048</td>
<td>20.2279</td>
<td>13.1398</td>
</tr>
<tr>
<td>(H_0: r = 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unit root test on the net-of-seigniorage ($\sigma^{OC}$) surplus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADF$_{SRC}$</td>
<td>-1.3319</td>
<td>-2.5710</td>
<td>-4.6976</td>
<td>-1.3269</td>
<td>-2.1622</td>
<td>-3.1871</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td><strong>Cointegration test between total expenditures and net-of-seigniorage ($\sigma^{OC}$) revenues</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\lambda_{MAX}) test</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(H_0: r = 0)</td>
<td>5.3073</td>
<td>11.7377</td>
<td>19.8559</td>
<td>12.3876</td>
<td>17.4073</td>
<td>12.5672</td>
</tr>
<tr>
<td>(r \leq 1)</td>
<td>3.4225</td>
<td>5.3890</td>
<td>3.4225</td>
<td>2.6586</td>
<td>2.9373</td>
<td>2.3331</td>
</tr>
<tr>
<td><strong>Trace test</strong></td>
<td>8.7099</td>
<td>17.1266</td>
<td>23.2784</td>
<td>15.0463</td>
<td>20.3447</td>
<td>14.9003</td>
</tr>
<tr>
<td>(H_0: r = 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the ADF$_{SRC}$ the lag length is chosen according to the 'Schwarz Bayesian Criterion'. In parenthesis is the number of lagged differenced terms used in the regression. Critical values of the DF$_{SRC}$ test (MacKinnon, 1991): -2.60(10%), -2.93(5%) and -3.58(1%). Critical levels of the Johansen test (Pesaran, Shin and Smith, 1997): 15.87/20.18 (5%) and 13.81/17.88 (10%) for the \(\lambda_{MAX}/Trace\) tests and the null hypothesis of \(r=0\), and 9.16 (5%) and 7.53 (10%), for both tests of \(r\leq 1\).

Whatever the underlying measure of seigniorage adopted, the results basically confirm all the main conclusions previously obtained in chapter 4.\(^\text{13}\) Only Germany and the Netherlands follow sustainable policies, and only the former presents a unitary cointegrating coefficient, suggesting a bounded debt-GDP ratio. The seigniorage revenues do not seem to have played a decisive role in determining whether the fiscal

\(^{13}\) In Table 5.2, the rejection of no cointegration in France is not statistically meaningful since the cointegrating coefficient is not significant. The (very weak) rejection of a unit root in the series of the net-
policies are sustainable, which is consistent with the relatively negligible magnitude of these revenues in most countries.

However, this result does not rule out the possibility that governments have used in fact this source of revenues to finance their expenditures. Given that monetization may ultimately cause higher inflation, it is important to evaluate whether in the past the level of the deficit has influenced the total of seigniorage revenues collected and, ultimately, the rate of inflation. This is the objective of the next empirical sections in this chapter.

5.5. The relation between deficits and seigniorage or inflation

The main purpose of this section is to examine whether unsustainable fiscal policies have constituted a constraint to monetary policy, leading the monetary authorities to allow money creation and higher inflation in order to ease the debt burden. The section is divided in two parts, and investigates the relationship between the central government’s budget deficit and the time series of, respectively, seigniorage and inflation. The first part provides an indirect approach to test the effects of the deficit on inflation, concentrating on the particular case of monetization, while the second part allows the consideration of all channels linking directly the fiscal deficits to the price level.\textsuperscript{14}

A VAR framework is adopted, allowing feedback effects and avoiding the arbitrary choice of the endogenous variables. The univariate statistical properties of the series involved were analysed with different unit root tests. Since the results are basically equivalent, Table 5.3 displays only the results of the ADF\textsubscript{ST} test, where the number of lagged dependent variables is chosen according to a sequential testing procedure defined in section 3.1.2. Any conflicting result will be reported below. The regressions do not

\footnote{\textsuperscript{14}Hondroyiannis and Papapetrou (1997) report evidence for Greece that deficits have an indirect effect on inflation through money growth (as tested in the previous section) rather than a direct effect (tested in this section).}
include a time trend since this deterministic variable is not significant. The inflation rate, $\pi^{\text{CPI}}$, was computed as the growth rate of the consumer price index.

The hypothesis of a unit root in the series of seigniorage revenues is rejected in all countries, except the UK, when the 'cash-flow' definition is employed, but only in Germany and the Netherlands when the 'opportunity cost' definition is used. This is not surprising, given the way both measures are constructed. The inflation rate seems to be stationary in Belgium, France and perhaps also in Germany, given the proximity to the 5% significance level. The total surplus has already been widely confirmed to be a stationary process only in Germany.

This disparity of results makes it difficult to examine the relation between deficits and seigniorage or inflation with a homogeneous, directly comparable methodology. Different procedures have to be employed, according to the stochastic characteristics of the underlying time series. With only stationary variables, the standard OLS regression and block causality tests can be applied. Conversely, if all variables involved are I(1), the cointegration procedures will be employed. Finally, when the model includes a mixture

---

**Table 5.3: ADF$_{ST}$ unit root tests on the series of seigniorage and inflation**

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^{CF}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3.9653</td>
<td>-4.6878</td>
<td>-4.0826</td>
<td>-3.7633</td>
<td>-4.6255</td>
<td>-2.3051</td>
</tr>
<tr>
<td>(0)</td>
<td>(0)</td>
<td>(2)</td>
<td>(0)</td>
<td>(2)</td>
<td>(2)</td>
<td>(2)</td>
</tr>
<tr>
<td>$\sigma^{OC}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0564</td>
<td>-1.3439</td>
<td>-3.7985</td>
<td>-1.6886</td>
<td>-4.1840</td>
<td>-1.1331</td>
</tr>
<tr>
<td>(2)</td>
<td>(0)</td>
<td>(2)</td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
<td>(0)</td>
</tr>
<tr>
<td>$\pi^{CPI}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0)</td>
<td>(0)</td>
<td>(2)</td>
<td>(0)</td>
<td>(3)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
</tbody>
</table>

In the ADF$_{ST}$ test the lag length is chosen according to the 'Sequential Testing' procedure. In parenthesis is the number of lagged differenced terms used in the regression. The asterisks (*), (**), and (***) indicate rejection of the null hypothesis of nonstationarity at the 10, 5 and 1 per cent significance levels, respectively. Critical values (MacKinnon, 1991): -2.60(10%), -2.93(5%) and -3.58(1%).
of stationary and nonstationary variables, the analysis follows the ARDL 'bounds procedure' suggested in Pesaran, Shin and Smith (1999). Being a recent approach, and not presented earlier in section 3.1.2, it is convenient to give a general view of the bounds-ARDL procedure used in this section.

The 'bounds testing approach' of Pesaran, Shin and Smith (1999), henceforth PSS, presents the major advantage of being adequate to test the existence of a long-run level relationship whether the underlying regressors are I(0) or, at maximum, I(1). It avoids the pre-testing problems involved in the unit root tests. The appropriate testing procedure depends on the specification of the deterministic variables included in the model to be estimated. For example, if the underlying level variables are allowed to follow a linear trend under the null hypothesis, the model with unrestricted intercepts and no trends (model III in PSS) is the more adequate. This model is based on the ECM specification

\[
\Delta y_t = \alpha + \Pi_y y_{t-1} + \Pi_x x_{t-1} + \sum_{j=1}^{q-1} \theta_j \Delta y_{t-j} + \sum_{j=1}^{q-1} \phi_j \Delta x_{t-j} + \varepsilon_t,
\]

where \( y_t \) represents the dependent variable, \( x_t \) is a vector (a scalar vector if the test involves only two variables as in the application below) of possibly long-run forcing variables, \( \Pi_y \) and \( \Pi_x \) are the respective long-run multipliers, and \( \theta_j \) and \( \phi_j \) represent the short-run dynamic coefficients. PSS include in the right-hand side of equation (5.16) the current first-differenced \( x_t \) variables. However, Pesaran and Pesaran (1997) argue that they should not be included if there is no a priori knowledge of which is (are) the long-run forcing variable(s). In this case of uncertainty, regression (5.16) is estimated for all variables, and the results of the tests compared to identify the direction of the long-run relationship. This seems the more adequate procedure in terms of the present study. Although the main purpose is to examine whether the fiscal problems of a country have an effect on monetization and inflation, it is advisable not to disregard the possibility of opposite effects, as noted above.

\[\text{18} \text{ A first, unpublished version of this paper (Pesaran, Shin and Smith, 1996) only covered models with unrestricted deterministic terms. It did not consider for example model (5.17) below, which ensures that the deterministic behaviour of the level variables is invariant to the result of the test, i.e., whether the null hypothesis is rejected or not.}\]

\[\text{19} \text{ Their classification follows the ordering of the models considered for example in Johansen (1994).}\]
If the variables are not supposed to follow a deterministic trend, but a constant is allowed in the error correction term, the test is based on a model with a restricted intercept and no trend (model II in PSS)

\[ \Delta y_t = \Pi_y(y_{t-1} - \alpha_y) + \Pi_x(x_{t-1} - \alpha_x) + \sum_{j=1}^{p} \theta_j \Delta y_{t-j} + \sum_{j=1}^{q} \phi_j \Delta x_{t-j} + \varepsilon_t. \]

The choice of the correct lag structure of the model is also a very important issue in these tests, although no particular strategy has been suggested so far to be the most efficient. A trade-off has to be considered, between the problems of overparameterization in small samples and the necessity to avoid serial correlation in the residuals. Two alternative strategies will be followed here. The first allows a very flexible choice for the dynamic lag structure of the model, by choosing a lag order \( p/q \) so that the element of order \( p+1/q+1 \) is rejected at the 10% significance level. The second imposes an order \( p=q \) so that the joint significance of the \( p+1=q+1 \) terms is not rejected with an \( F \)-test. In the particular case of the tests performed here, both strategies lead to qualitatively similar results in almost all cases, as shown in the tables below.

When model III, equation (5.16), is considered, the test is based on the standard \( F \)-statistic (or Wald statistic) of the joint significance of the lagged levels of all the variables included. With model II, the joint test is extended to the significance of the intercept term. The null hypothesis is that there is no long-term relationship between the variables considered. In general terms,

\[ H_0: \Pi_y=0 \text{ and } \Pi_x=0, \]

\[ H_a: \Pi_y \neq 0 \text{ or } \Pi_x \neq 0. \]

The asymptotic distribution of these \( F \)-test statistics is non-standard under the null, but it is derived and tabulated in PSS (see their Tables C1). Given the uncertainty about the order of integration of the variables, there are two sets of critical values, each assuming a

---

20 PSS also suggest an alternative test, based on the \( t \)-statistic of the significance of the lagged dependent variable, which basically constitutes an extension of the cointegration test proposed by Banerjee, Dolado and Mestre (1998) to the case where the order of integration of the variables is uncertain.

21 It is important to note that this allows the theoretical possibility of a 'degenerate' long-run relationship between the variables, under the null hypothesis, in which the lagged levels of the forcing variables do not enter the equation for \( \Delta y \). This possibility will be carefully considered in the empirical applications.
polar case of all variables being I(0) or all I(1). If the test statistic is above the band formed by the 'critical values bounds', it suggests a rejection of the null hypothesis of no long-term relationship. Conversely, if it falls below the band, the null cannot be rejected. Otherwise, if the statistic falls inside the band, no conclusive inference can be drawn, and further information about the correct order of integration of the variables is required.

However, according to assumption 3 of PSS, this bounds testing approach is not valid when the null hypothesis is rejected in more than one of the first-differenced regressions. In terms of equations (5.16) and (5.17), this entails the identification requirement that the lagged level of $y_t$ cannot enter any of the equations for the $x_t$ variables (although there may still be Granger-causality from $y_t$ to $x_t$ in the short run, a possibility which will be explored below).

After concluding in favour of the existence of a long-run relationship between the variables (and its direction), the second stage of the procedure is to use the two-step autoregressive distributed lag (ARDL) approach suggested in Pesaran and Shin (1999) to examine the parameters of that relationship. This approach yields consistent estimates of the long-run coefficients regardless of whether the underlying regressors are I(0) or I(1). The order of the ARDL is selected according to the Schwarz Bayesian Criterion (SBC), since this was found by the authors to perform slightly better than the Akaike Information Criterion (AIC). A maximum lag length of four was allowed in the tests below.

5.5.1. Deficits and seigniorage

Table 5.4 presents the summary results of the application of the ARDL bounds testing procedure to the relationship between the total surplus and the amount of seigniorage, measured by the cash-flow definition. The upper part of the table applies the bounds procedure to all countries, while the lower part presents the estimates of the model in those cases where a long-run significant relationship was found.

In exactly half the countries in the sample there is indication of a long-run relationship between total deficits and seigniorage (cash-flow definition), with the former being the forcing-variable. However, contrary to the suggested by the standard theory, this relation is negative in all three cases. Furthermore, in Italy and the Netherlands the results in the lower part of the table suggest a 'degenerate' long-run relationship between the levels of
the variables. In this case there are no long-run effects running from the deficit to the seigniorage revenues. The long-run relationship suggested in the upper part of the table involves only $\sigma_{t,F}$ and the deterministic term.

Table 5.4: Bounds-ARDL test procedure applied to the relation $ts_t, \sigma_{t,F}$

<table>
<thead>
<tr>
<th>dependent variable</th>
<th>order chosen with t-tests</th>
<th>order chosen with F-tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p  q  III</td>
<td>II</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>1  4</td>
<td>9.8480**</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>1  2</td>
<td>2.4960</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>4  4</td>
<td>3.4718</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>4  1</td>
<td>5.0266</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>3  2</td>
<td>10.6943**</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>2  1</td>
<td>1.0805</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>3  1</td>
<td>2.0463</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>1  4</td>
<td>3.1398</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>3  1</td>
<td>14.3145**</td>
</tr>
<tr>
<td>$\Delta \sigma_{t,F}$</td>
<td>3  3</td>
<td>4.0236</td>
</tr>
</tbody>
</table>

Country  Depend. variable  underlying ARDL order (p,q)  R²  regressor  intercept  ECT coefficient
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>$\sigma_{t,F}$</td>
<td>(2,1)</td>
<td>33.3%</td>
<td>0.0697 (0.0334)</td>
<td>0.0090 (0.0021)</td>
</tr>
<tr>
<td>Italy</td>
<td>$\sigma_{t,F}$</td>
<td>(2,2)</td>
<td>45.9%</td>
<td>0.0163 (0.0769)</td>
<td>0.0174 (0.0065)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>$\sigma_{t,F}$</td>
<td>(1,1)</td>
<td>1.7%</td>
<td>0.0341 (0.0404)</td>
<td>0.0064 (0.0013)</td>
</tr>
</tbody>
</table>

Upper part: The asterisks (**) indicate rejection of the null hypothesis of no long-run relationship at the 5% significance level. The question mark (?) reveals that the result is not conclusive, falling inside the band. All other results indicate that it is not possible to reject the null, since the statistic falls below the band. The hyphen (-) shows that the chosen order is too low to allow the test. Critical values bounds (Pesaran, Shin and Smith, 1999): [5.43/6-24] for model II and [4.94/5.73] for model III, for 5% significance levels. Lower part: order chosen by the SBC, starting from a maximum order of 4. Below the estimates are the standard errors, in parenthesis.
In the case of Belgium, the only country where a significant long-run relationship was found, one possible explanation for the negative relation between deficits and seigniorage may simply be the influence of the business cycle. A period of economic growth is usually associated with lower deficits (tax revenues increase while transfers diminish), and higher seigniorage revenues, through the positive influence of the inflation and growth rates and the increased demand for real money balances, as noted in section 5.3. Another hypothesis is that the monetary authorities react to expansionary fiscal policies with restrictive monetary policies, in order to avoid a rising inflation rate. The result of this reaction is a worsening of the fiscal situation. Besides the loss of the seigniorage revenues, a restrictive monetary policy leads to higher interest rate and therefore a higher interest burden of the debt, a significant problem in the case of high debt countries like Belgium.

Table 5.5 presents the results of applying the above methodology to the case of the 'opportunity cost' measure of seigniorage. Again, in two of the three cases where the bounds testing procedure suggested a long-run relationship between the levels of the total deficit and the seigniorage revenues ('opportunity cost' measure), this relation seems to be 'degenerate'. Therefore, the only significant relationship arising from Table 5.5 concerns Italy. Here it was found that the seigniorage revenues have a long-run positive impact on the level of the Italian total deficit. The reason for the positive sign of the relation lies possibly on the influence of the interest rate. Given this definition of seigniorage, these revenues increase with the interest rate which, at the same time, expand the interest payments on the debt, possibly more than proportionally, especially in high debt countries like Italy. The next chapter will consider in more detail the influence of the interest rate in the fiscal dynamics.
Table 5.5: Bounds-ARDL test procedure applied to the relation $\Delta s_t$, $\sigma_{t}^{OC}$

<table>
<thead>
<tr>
<th>Country</th>
<th>Dependent variable</th>
<th>underlying ARDL</th>
<th>long-run coefficients</th>
<th>ECT coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>order (p,q)</td>
<td>$R^2$</td>
<td>regressor</td>
<td>intercept</td>
</tr>
<tr>
<td>Germany</td>
<td>$\sigma_{t}^{OC}$</td>
<td>(2,1)</td>
<td>49.8%</td>
<td>0.0724</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0082</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.2390</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0793)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.1198)</td>
</tr>
<tr>
<td>Italy</td>
<td>$ts$</td>
<td>(3,1)</td>
<td>90.6%</td>
<td>-6.6983</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0289</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.2515</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.3651)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0226)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0680)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0094)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.1364)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>$\sigma_{t}^{OC}$</td>
<td>(2,1)</td>
<td>33.5%</td>
<td>-0.0108</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0059</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.4632</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0094)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.1364)</td>
</tr>
</tbody>
</table>

Upper part: The asterisks (**) indicate rejection of the null hypothesis of no long-run relationship at the 5% significance level. The question mark (?) reveals that the result is not conclusive, falling inside the band. All other results indicate that it is not possible to reject the null, since the statistic falls below the band. The hyphen (-) shows that the chosen order is too low to allow the test. Critical values bounds (Pesaran, Shin and Smith, 1999): [5.43/6-24] for model II and [4.94/5.73] for model III, for 5% significance levels. Lower part: order chosen by the SBC, starting from a maximum of order 4. Below the estimates are the standard errors, in parenthesis.

For some of the countries in the sample, it is possible to examine the long-run relationship between the seigniorage revenues and the deficits by using the more standard VECM procedure. This can only be applied to the 'opportunity cost' definition, given the stationary properties of the alternative measure. Table 5.6 presents the results of the Johansen cointegration tests on the series of seigniorage and total surplus. The order of the underlying VAR was chosen according to the Schwarz Bayesian Criterion (SBC), and the model includes a restricted intercept and no trend.
Table 5.6: Johansen cointegration tests on seigniorage ($\sigma^{OC}$) and total surplus ($ts_t$)

<table>
<thead>
<tr>
<th>Order</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{MAX}$ test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0: r=0$</td>
<td>7.0842</td>
<td>12.7643</td>
<td>10.2648</td>
<td>12.1875</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0: r\leq1$</td>
<td>5.8336</td>
<td>2.5555</td>
<td>1.4581</td>
<td>1.5744</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trace test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0: r=0$</td>
<td>12.9178</td>
<td>15.3198</td>
<td>11.7229</td>
<td>13.7619</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Critical values (Pesaran, Shin and Smith, 1997): 15.87/20.18 (5%) and 13.81/17.88 (10%) for the $\lambda_{MAX}$/Trace tests and the null hypothesis of $r=0$, and 9.16 (5%) and 7.53 (10%) for both tests of $r\leq1$.

No long-run equilibrium relationship could be found in any of the four countries tested in the table. In the particular case of Germany both variables seem to be stationary, which was confirmed by the fact that applying the cointegration test of Table 5.6 leads to a rejection of both null hypotheses. In this situation, a simple OLS regression is adequate, producing the estimates (standard errors in parenthesis)

$$ts_t = -0.3445 \sigma^{OC} - 0.0089,$$

$$\text{(1.1617)} \quad \text{(0.0090)}$$

confirming the non-existence of a significant relationship between the two variables.

In spite of this lack of evidence of a long run relationship, there is still the possibility of finding some causal effects between the variables, arising from the short-run dynamics. Table 5.7 displays the results of the non-causality tests. Given the stochastic characteristics of the variables, and the fact that no cointegrating relationship was found between them, the LA-VAR test seems to be the most adequate procedure.

The only significant causal effects run from the seigniorage revenues, measured by the opportunity cost, to the total surplus in Belgium and Italy. Again, this seems to reflect merely the importance of the interest rate in the budget dynamics of the high debt countries. A rise in the interest rate increase this measure of the seigniorage revenues, but will also expand the cost of the debt and henceforth the deficit.

The evidence that the connection between this measure of seigniorage and the total deficit results from the interest rate, rather than from changes in the monetary base, is grounded on two further observations: if the opportunity cost definition of seigniorage is substituted by the cash-flow definition (not directly affected by interest rate movements),
no evidence of causality remains; on the other hand, in Italy for example, although the
same happens in all other countries, there is no evidence of a causal relationship between
the monetary base (m) and the total surplus (as ratios of GDP), but there is a strong
evidence of a causal relation running from the nominal (and real) interest rate to the total
surplus.

Table 5.7: LA-VAR non-causality tests - seigniorage (σ^CF and σ^OC) and total surplus (ts)

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>'cash-flow' measure of seigniorage (σ^CF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>σ^CF → ts</td>
<td>1.5134</td>
<td>3.8551 *</td>
<td>2.3811</td>
<td>0.0720</td>
<td>0.1423</td>
<td>0.0016</td>
</tr>
<tr>
<td>ts → σ^CF</td>
<td>1.8173</td>
<td>0.0840</td>
<td>0.1742</td>
<td>0.0467</td>
<td>0.6884</td>
<td>0.5539</td>
</tr>
<tr>
<td>order</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>'opportunity cost' measure of seigniorage (σ^OC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>σ^OC → ts</td>
<td>14.7043 ***</td>
<td>3.4520 *</td>
<td>0.2962</td>
<td>11.8910 ***</td>
<td>2.3724</td>
<td>0.2385</td>
</tr>
<tr>
<td>ts → σ^OC</td>
<td>0.1423</td>
<td>0.6247</td>
<td>0.4339</td>
<td>4.1347</td>
<td>0.0025</td>
<td>0.2207</td>
</tr>
<tr>
<td>order</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

† A point dummy (1975) was added to the model to solve problems of non-normality. The asterisks (***)
and (*) indicate rejection of the null hypothesis of no causality at the 1% and 10% levels of significance,
respectively. In brackets are the p-values of the test (χ² distribution).

In sum, this section has shown that for the countries and the sample period analysed,
there is no evidence that deficits tend to be monetized. The evolution of the deficit does
not seem to be directly related to a resort to seigniorage revenues. The only significant
relation found in this section indicates that increasing levels of the 'opportunity cost'
definition of seigniorage originate higher budget deficits. However, this may simply
reflect the importance of the interest rate in the budget dynamics. Chapter 6 will
investigate this hypothesis further.
5.5.2. Deficits and inflation

Although the results of the previous section have shown that the central government deficits do not lead to an increasing recourse to seigniorage revenues, there are other channels besides central bank monetization through which the budget could cause, or benefit, from higher levels of inflation. In addition to the already mentioned inflation-tax on the stock of debt, and the nominal features of the tax/transfer system (see the discussion in section 5.2), two other channels have been referred in the literature, operating respectively on the demand-side and on the supply-side. The first channel of transmission works through the ‘wealth effect’ (see Dwyer, 1982, for example). An increase in the stock of public bonds held by the public increases the perceived private wealth (provided the Ricardian Equivalence Theorem does not hold), promotes higher spending and therefore inflation. The second effect is through the cost of the production factors (Wray, 1997). An expenditure-driven increase in the deficit may lead to a relative scarcity of some economic resources, increasing their cost and therefore the price of the final products.

This section investigates the relation between the levels of deficits and inflation, using the same methodology followed above. As before, the methods used here are more suitable to analyse relations between variables showing distinct orders of integration. Table 5.8 presents the results of the bounds testing procedure and the ARDL modelisation of the long-run relationship between the series of total surplus and inflation. In two cases it was necessary to include dummies in the model to solve problems of non-normality. The asymptotic critical values are not affected by the inclusion of these dummies (Pesaran, Shin and Smith, 1999).

The results in Table 5.8 suggest the existence of a long-run relationship between inflation and the total deficit in three countries, positive in the cases of Belgium and Italy, negative in France. This latter result is difficult to explain, but once again demonstrates the idiosyncratic behaviour of the French fiscal variables relatively to the rest of the EU. The period of high inflation, in the second half of the seventies, corresponds in France with the period of lower deficits. Since the early eighties the drop in the inflation rate was
accompanied by a growing budget deficit. This inverse relationship may simply reflect the influence of the business cycle.

Table 5.8: Bounds-ARDL test procedure applied to the relation $t_{St}, \pi_{t} CPI$

<table>
<thead>
<tr>
<th>dependent variable</th>
<th>order chosen with $t$-tests</th>
<th>Order chosen with $F$-tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>order</td>
<td>model</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>q</td>
</tr>
<tr>
<td>B $\Delta t_{St}$</td>
<td>1 1</td>
<td>6.9379**</td>
</tr>
<tr>
<td></td>
<td>1 1</td>
<td>5.0971'</td>
</tr>
<tr>
<td>$\Delta \pi CPI$</td>
<td>1 2</td>
<td>1.5435</td>
</tr>
<tr>
<td>F $\Delta t_{St}$</td>
<td>1 3</td>
<td>11.5857**</td>
</tr>
<tr>
<td>$\Delta \pi CPI$</td>
<td>4 1</td>
<td>12.0395**</td>
</tr>
<tr>
<td>G $\Delta t_{St}$</td>
<td>3 2</td>
<td>5.4228'</td>
</tr>
<tr>
<td>$\Delta \pi CPI$</td>
<td>2 1</td>
<td>10.0335**</td>
</tr>
<tr>
<td>I $\Delta t_{St}$</td>
<td>1 3</td>
<td>1.6896</td>
</tr>
<tr>
<td>$\Delta \pi CPI$</td>
<td>1 1</td>
<td>2.7826</td>
</tr>
<tr>
<td>N $\Delta t_{St}$</td>
<td>4 1</td>
<td>3.7730</td>
</tr>
<tr>
<td>$\Delta \pi CPI$</td>
<td>1 1</td>
<td>4.3743</td>
</tr>
<tr>
<td>UK $\Delta t_{St}$</td>
<td>1 1</td>
<td>2.8832</td>
</tr>
</tbody>
</table>

$\Delta \pi CPI$

<table>
<thead>
<tr>
<th>Country</th>
<th>depend. variable</th>
<th>underlying ARDL order (p,q)</th>
<th>R²</th>
<th>long-run coefficients regressor</th>
<th>intercept</th>
<th>ECT coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>$t_{St}$</td>
<td>(2,1)</td>
<td>93.7%</td>
<td>-1.4379 (0.8095)</td>
<td>0.0235 (0.0428)</td>
<td>-0.0887 (0.0418)</td>
</tr>
<tr>
<td>France</td>
<td>$\pi_{t} CPI$</td>
<td>(2,2)</td>
<td>72.9%</td>
<td>2.2728 (0.5631)</td>
<td>0.0912 (0.0125)</td>
<td>-0.4578 (0.0873)</td>
</tr>
<tr>
<td>Italy</td>
<td>$t_{St}$</td>
<td>(3,2)</td>
<td>91.4%</td>
<td>-0.7907 (0.1918)</td>
<td>-0.0203 (0.0162)</td>
<td>-0.2139 (0.0591)</td>
</tr>
</tbody>
</table>

Upper part: The asterisks (**) indicate rejection of the null hypothesis of no long-run relationship at the 5% significance level. The question mark (?) reveals that the result is not conclusive, falling inside the band. All other results indicate that it is not possible to reject the null, since the statistic falls below the band. The hyphen (-) shows that the chosen order is too low to allow the test. Critical values bounds (Pesaran, Shin and Smith, 1999): [5.43/6-24] for model II and [4.94/5.73] for model III, for 5% significance levels. Lower part: order chosen by the SBC, starting higher from a maximum order of 4. Below the estimates are the standard errors, in parenthesis. 1 the model includes point dummy variables in 1958 and 1981 to solve problems of non-normality. 2 the model includes a point dummy variable in 1958 to solve problems of non-normality.

Smithin (1994) also reports evidence of a negative correlation between deficits and inflation in the US, while King (1998) finds a negative relation between debt and inflation in a pooled group of OECD countries.
In the cases of Belgium and Italy, the relation between the deficit and the inflation rate is positive, with the latter being the long-run 'forcing variable'. This contradicts the conventional economic theories in two aspects. First, it is suggested that, predominantly, deficits have an effect on the inflation rate, not the other way round. Second, the effects of inflation on the deficit should be negative. The budget benefits from the inflation-tax on the monetary base and on the debt, and from the nominal features of the tax system.

Further tests are performed below to check the robustness of these results. For Italy, the Netherlands and the UK, it is possible to complement the ARDL procedure with the more conventional VECM, since both variables involved are I(1). Table 5.9 reports the results of the Johansen cointegration tests for these countries.

Table 5.9: Johansen cointegration tests - total surplus ($t_s$) and inflation ($\pi_t^{CPI}$)

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{MAX}$ test</td>
<td>$Trace$ test</td>
<td>$\lambda_{MAX}$ test</td>
<td>$Trace$ test</td>
</tr>
<tr>
<td>order</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$H_0$: r=0</td>
<td>18.0289**</td>
<td>20.8593**</td>
<td>12.4690</td>
</tr>
<tr>
<td>$H_0$: r\leq1</td>
<td>2.8304</td>
<td>2.8304</td>
<td>4.1417</td>
</tr>
</tbody>
</table>

Critical values (Pesaran, Shin and Smith, 1997): 15.87/20.18 (5%) and 13.81/17.88 (10%) for the $\lambda_{MAX}$/$Trace$ tests and the null hypothesis of $r=0$, and 9.16 (5%) and 7.53 (10%) for both tests of $r\leq1$.

The null hypothesis of no cointegration could only be clearly rejected in the case of Italy.

The cointegrating vector, normalised for the $t_s$ variable (standard errors in parenthesis), is

$$t_s = -0.8714 \pi_t^{CPI} - 0.0116, \quad (0.2215) \quad (0.0185)$$

again revealing the existence of a positive long-run relationship between the deficits and inflation. Furthermore, the error correction term was found to be significant only in the equation for $t_s$, confirming all the findings of the ARDL procedure. For the other two countries analysed, the previous results of no long-run relationship were also confirmed by the VECM procedure. In the case of Germany, where both variables are stationary, estimation by OLS provided the equation.
\[ t_s = -0.0125 \, \pi_{CPI} - 0.0112, \]
\[ (0.0635) \quad (0.0024) \]

again showing a negative, but not statistically significant relation. Further complementary analysis may be provided by the non-causality tests presented in Table 5.10.

Table 5.10: LA-VAR non-causality tests - inflation (\( \pi_{CPI} \)) and total surplus (\( t_s \))

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_{CPI} \rightarrow t_s )</td>
<td>7.4837***</td>
<td>5.5849**</td>
<td>33.8739***</td>
<td>11.6704***</td>
<td>3.1387*</td>
<td>0.7193</td>
</tr>
<tr>
<td>( t_s \rightarrow \pi_{CPI} )</td>
<td>1.8058</td>
<td>3.1387*</td>
<td>3.3428</td>
<td>0.6101</td>
<td>0.7582</td>
<td>0.0022</td>
</tr>
<tr>
<td>order</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

The asterisks (***) (***) and (*) indicate rejection of the null hypothesis of no causality at the 1%, 5% and 10% levels of significance, respectively. In brackets are the p-values of the test (\( \chi^2 \) distribution).

For most countries in the sample, there is a strong causal relationship running from the inflation rate to the total deficit. The opposite effect was also found in the case of France, confirming the findings of the ARDL procedure, but it is significant only at the 10% level.

In sum, this section has shown that deficits do not seem to contribute to a general price increase. Overall, during the sample period considered, the European governments did not resort to inflation to solve their fiscal problems. On the contrary, the findings of this section suggest that inflation has harmful effects on the budget. If the economic agents expect high inflation rates, they incorporate this in the nominal interest rate demanded on public bonds, increasing the debt service costs of the government and therefore the deficits. This transmission channel was advanced by Barro (1979) and empirically confirmed for example in Dwyer (1982) and Joines (1985). It shows that a positive relation between deficits and inflation can be found even outside a fiscal dominance regime, and strengthens the case for a careful examination of the direction of causality.
5.6. A public finance model of optimal revenue collection

An alternative to the VAR approach used in the previous section, is to adopt an intertemporal public finance optimization model of the choice between different sources of government revenues (taxes and seigniorage). This model may help clarify whether monetary base growth and inflation have been determined by revenue needs in the past and whether this is likely to happen in the future.

The theoretical models of optimal seigniorage can be traced back to the works of Bailey (1956), Cagan (1956) and Friedman (1969). However, these studies did not consider explicitly the consequences of this seigniorage policy on the government’s budget constraint, being more concerned with aspects of revenue maximization. It was Phelps (1973) who included this question of optimal seigniorage collection in a more general public finance perspective of optimal taxation. Many more studies followed: Calvo (1978), Drazen (1979), Helpman and Sadka (1979), Kimbrough (1986), and Lucas (1986), among others.

At the same time, Barro (1979) disregards the monetary source of revenues and presents the ‘tax-smoothing’ hypothesis, according to which the marginal costs of (distortionary) taxation should be equalized intertemporally. The tax revenues should be set to pay for the permanent component of expenditures, the temporary changes of spending being covered by issuing/paying debt. As a result of the theory, changes in the tax rate should be unpredictable.

Mankiw (1987) incorporates the views of Phelps in Barro’s work, to analyse empirically the joint effects of the fiscal and monetary sources of government receipts. He argues that an optimizing government should choose the value of the seigniorage revenues compatible with the equalization of the deadweight loss from this and other sources of income, and derives the testable implications of the model: the inflation rate and the average tax rate should be positively correlated. Applying the test to US data, he finds evidence in favour of the model.

21 This type of model provides the formal justification for the neoclassical public finance argument for different inflation rates across countries.
22 He would later (Barro, 1983) introduce seigniorage in his model of optimal taxation.
Grilli (1989) extends Mankiw's model of public finance to the EU case, and establishes different testable implications. He derives the time series properties of the seigniorage and tax rates and concludes that optimality in the choice of the revenues requires a cointegrating relationship between these two variables. Applying this test to ten EU countries, he presents evidence that the model is in fact followed at least in half of these countries.

These discrepancies between the behaviour of different governments within the EU could have important consequences in EMU, and emphasise the importance of the exchange rate regime on the monetization decisions. In a fixed exchange rate regime the governments lose the ability to set independently the preferred rate of inflation, according to their financing needs. At the same time, the EMU imposes important constraints on the autonomous collection of alternative sources of revenues.

In fact, since the possibility of EMU began to be seriously considered in the late eighties, a number of authors have insinuated that seigniorage might be a significant source of government revenues in some European Union countries, with important consequences in terms of monetary integration. Dornbusch (1988), for example, claims that this source of revenues may be appropriate in countries with inefficient tax systems and/or high debt-GDP ratios. He draws attention to the possibility that the loss of these revenues in some EU members may be an excessive cost of participation in the EMU. More recently, Buiter (1995a) argues that this may be too extreme, but still maintains that countries with high debt to GDP ratios should have the option of a burst of unanticipated inflation to devalue public debt just before joining EMU.

5.6.1. The model

The main rationale behind the model of optimal revenue collection is that both ways of financing the government expenditures (taxes and seigniorage) involve efficiency costs. Taxation may have negative effects for example on the decisions concerning labour supply, savings and investment. Monetization, and inflation, has negative effects in the

---

demand for cash balances, distributional consequences (redistributing wealth between debtors and creditors), and leads to inefficiencies because of the existence of nominal contracts.

The objective of the authorities is therefore to minimize the costs of obtaining the required amount of revenues. In the optimal point, efficiency requires that the marginal deadweight cost from each revenue source will be equalized, intra- and intertemporally, and more revenues will be collected from the less costly source. Formally, and following Grilli (1989), Poterba and Rotemberg (1990) and Trehan and Walsh (1990), the intertemporal, discounted, cost function to be minimized can be represented by the following generalized, constant elasticities, CES welfare function

\[ \sum_{j=1}^{n} E_t \delta_{t,j} \left\{ \left[ \frac{1}{1+\alpha} \right] \frac{(1+\pi_{t+j})^{-1}}{1-\beta} \right\}, \]

where \( \delta_{t,j} = \prod_{i=1}^{j} (1+r_{t+i})^{-1} \) is the intertemporal discount factor, \( r_t \) is the real interest rate, \( t_t \) represents, as before, the average tax rate, \( \pi_t \) is the rate of inflation and \( \alpha \) and \( \beta \) (both positive) are constant elasticities (dependent on the government’s preferences). Adopting the suggestion of Trehan and Walsh (1990), the objective function includes the stochastic terms \( \phi_t \) and \( \theta_t \) that allow for changes in the distortionary cost functions of, respectively, the tax and inflation rates. These changes in the cost functions may result, for example, from transformations in the labour market, the tax system or the economy’s transaction and tax collecting technology.

The cost function depends positively on the intertemporal set of tax rates and inflation. For algebraic convenience, instead of modelling directly the costs of inflation, the cost function includes the benefits of deflation, entered with a negative sign. When minimizing the cost function, the authorities must obey the intertemporal budget constraint, here expressed in real terms as

\[ \sum_{j=1}^{n} \delta_{t,j} E_t \left[ t_{t+j} Y_{t+j} + \left[ M_{t+j} - M_{t+j-1} (1+\pi_{t+j})^{-1} \right] - G_{t+j} - D_t \right] = 0, \]

\( ^{26} \) This implicitly presupposes that countries with more developed and efficient tax systems should turn less to seigniorage revenues. Buiter (1995a) presents some doubts about the robustness of this proposition.
where $Y_t$, $M_t$, $G_t$, and $D_t$ are the real values of, respectively, output, base money, expenditures and interest-bearing government debt. In square brackets are the seigniorage revenues, measured as the change in the real value of base money ('cash-flow' measure).

Assuming that the monetary and fiscal authorities coordinate their policies in order to minimize the costs of revenue collection, subject to the restriction of the intertemporal budget constraint, it is possible to construct the Lagrange function and derive the conditions of optimization. From the first order conditions, the optimal combination of revenues that minimizes the distortionary costs is given by,

$$t^* = \frac{Y_t \theta_t (1 + \pi_t) \theta_t (1 + \eta)}{\phi_t M_{t-1} (1 + \mu)}.$$

The parameters $\eta$ and $\mu$ represent, respectively, the elasticity of real output with respect to the tax rate and the interest elasticity of real base money. Real government expenditures were considered exogenous to the tax and inflation rates in this model, which seems consistent with most of the previous empirical results.

In addition, for intertemporal optimality, it is necessary that the marginal decline in the cost of raising revenues, resulting from an additional, present value unit increase of government revenues available (given by the usual $A$ parameter of the Lagrangian function), is the same over time. This implies that

$$E_t \frac{t^* \pi_{t+1} \theta_{t+1}}{Y_{t+1} (1 + \eta)} = \frac{t^* \phi_t}{Y_t (1 + \eta)}, \text{ and}$$

$$E_t \frac{(1 + \pi_{t+1}) \theta_{t+1}}{M_t (1 + \mu)} = \frac{(1 + \pi_t) \theta_t}{M_{t-1} (1 + \mu)}.$$

Equation (5.20) is the static first-order condition. Equations (5.21) and (5.22) are the intertemporal first-order conditions (Euler equations), implying that the tax and inflation

---

27 It may be argued that, in reality, the monetary authorities do not seem to have revenue considerations in mind when choosing the optimal monetary policy. However, Poterba and Rotemberg (1990) and Trehan and Walsh (1990) argue that when a high spending government increases taxes and debt finance, the central bank reacts by purchasing those bonds in order to prevent an increase in interest rates resulting from the pressure on financial markets, as the government competes with private borrowers for scarce loanable funds. This increases the base money and therefore seigniorage revenues, resulting in the outcome implied by the model.
rates are smoothed over time. Taking natural logarithms from equations (5.20)-(5.22) to obtain linear equations yields, respectively,

\[ \pi_t = \frac{\alpha}{\beta} \ln t + \frac{1}{\beta} [\ln (1 + \mu) - \ln (1 + \eta)] + \frac{1}{\beta} (\ln M_{t-1} - \ln Y_t) + \frac{1}{\beta} (\ln \phi_t - \ln \theta_t), \]

(5.23)

\[ E_l \ln t_{t+1} = \ln t_t + \frac{1}{\alpha} (\ln Y_{t+1} - \ln Y_t) - \frac{1}{\alpha} (\ln \phi_{t+1} - \ln \phi_t) \]

and

(5.24)

\[ E_l \pi_{t+1} = \pi_t + \frac{1}{\beta} (\ln M_t - \ln M_{t-1}) - \frac{1}{\beta} (\ln \theta_{t+1} - \ln \theta_t). \]

(5.25)

The rate of inflation, \( \pi_t \), has been approximated by \( \ln (1 + \pi_t) \). Equation (5.23) shows that inflation is negatively correlated with output growth. Since economic growth provides additional revenue from money creation for the government (see equation (5.5), above), the higher the rate of growth the lower will be the optimal inflation rate required.

Assuming a zero or constant growth rate of real output and real base money, rational expectations, and no shifts in the distortionary costs associated with both revenue sources (\( \phi \) and \( \theta \) are constant), 28 equations (5.24) and (5.25) reveal that the optimal condition requires the series of inflation and the logarithm of the tax rate to be nonstationary, or I(1).

Equation (5.23) shows that the inflation and the (log of) tax rate must move together in the same direction, in response to a change in the revenue demands. The relative value of the response of each revenue source depends on the difference between their tax bases and on the difference between the respective distortionary costs (third and fourth terms on the right-hand side, respectively). If these differences are stationary, given that the inflation rate and the logarithm of the tax rates are nonstationary, these must be cointegrated to satisfy equation (5.23). The cointegrating parameter \( \alpha/\beta \) is a measure of the relative cost of both revenue sources.

Mankiw (1987), Poterba and Rotemberg (1990), and Goff and Toma (1993) test the implications of the model by performing a simple OLS estimation of equation (5.23), and

---

28 Or, less restrictively, that all these growth rates are stationary. The null hypothesis of a unit root in the series of real output and real base money growth has been strongly rejected for all countries in the sample, based on ADF_{TT} tests (not shown).
checking whether the coefficient of the tax rate is positive and significant. However, the non consideration of unit roots may lead to spurious results and invalidate their conclusions. Grilli (1989) and Trehan and Walsh (1990) consider this possibility and test for cointegration between the series of inflation and tax rates. Mixed conclusions have emerged from all these studies, mostly concerned with the US case. Instead of studying the time series implications of the model, Click (1998) analysed its cross-section properties in a large number of countries for the post-Bretton Woods period. He concludes that the optimal taxation model explains up to 40% per cent of the cross-country variation in seigniorage.

One problem with most of the above studies is the assumption of a constant velocity of base money (given in equation (5.23) by the difference between the logs of real base money and real output), a variable therefore usually included in the intercept term. Poterba and Rotemberg (1990) and Goff and Toma (1993) explicitly include this variable in their regressions, but do not consider the possibility that it may be nonstationary.

In reality, the process of technological innovations in the financial markets, and the change in the policy of banking reserve requirements, indicates that this assumption of a constant, or even a stationary velocity of money may be too strong. If this time series is nonstationary in practice, non-cointegration between the tax and inflation rates is not the correct null hypothesis to examine. Instead, the condition of optimality requires that the inflation rate, the logarithm of the tax rate and the velocity of money are cointegrated. This is the condition tested in the following section.

5.6.2. Empirical application

As established above, the model requires, first, that both series of inflation and (log of) tax rates are nonstationary and, second, that they are cointegrated (possibly also with the base money ratio, the inverse of velocity) with a positive cointegrating parameter. The unit root tests on the series of inflation have already been performed above and are displayed in Table 5.3. The rejection of the null hypothesis of nonstationarity in the cases

---

29 With nonstationary variables, their OLS estimates of the coefficients may be inconsistent. The usual t-statistics of the coefficients do not have a limiting distribution and therefore cannot be used to test the significance of the regression parameters.
of Belgium and France immediately precludes further investigation, and suggests that the model does not hold in these countries. Table 5.11 presents the results of the ADF$_{ST}$ unit root tests on the logarithm of the tax rate and of the base money relative to GDP.\textsuperscript{30} When significant, a trend is included in the regression.

Table 5.11: ADF$_{ST}$ unit root tests on the (log of) tax rate and base money to GDP ratio

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ln t_t$</td>
<td>-0.8553</td>
<td>-3.4662*</td>
<td>-1.7399</td>
<td>-0.4010</td>
<td>-0.8508</td>
<td>-3.9001*</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(3)</td>
<td>(0)T</td>
</tr>
<tr>
<td>$ln m_t$</td>
<td>-2.4375</td>
<td>-2.2861</td>
<td>-3.9237*</td>
<td>-0.6962</td>
<td>-1.6011</td>
<td>-2.0444</td>
</tr>
<tr>
<td></td>
<td>(0)T</td>
<td>(3)T</td>
<td>(1)T</td>
<td>(2)T</td>
<td>(0)</td>
<td>(3)T</td>
</tr>
</tbody>
</table>

In the ADF$_{ST}$ test the lag length is chosen according to the 'Sequential Testing' procedure, starting with a maximum of three lags. In parenthesis is the number of lagged differenced terms used in the regression. A $T$ indicates that a (significant) trend was included in the model. The asterisks (*), (**), and (***) indicate rejection of the null hypothesis of nonstationarity at the 10, 5 and 1 per cent significance levels, respectively. Critical values for 46 observations (MacKinnon, 1991): -2.60(10%), -2.93(5%) and -3.58(1%) for the model with intercept but no trend and -3.1914 (10%), -3.5134 (5%) and -4.1695 (1%) for the model with intercept and trend.

The rejection of the null hypothesis of a unit root in the series of the tax rate in the UK suggests immediately that this country does not follow the optimisation model, joining the group of Belgium and France. For the other countries, the first condition of nonstationarity of the series of inflation and tax rate is satisfied, delaying the conclusions to the analysis of cointegration. The presence of a unit root in most series of base money relative to GDP (the inverse of the velocity of money) suggests that all the previous studies disregarding this possibility may have provided inaccurate conclusions.

Table 5.12 displays the results of the Johansen cointegration tests. For Italy and the Netherlands, the model includes the three variables. In the particular case of Germany however, the velocity of money is stationary, and therefore the standard methodology of testing for cointegration between only the other two series is feasible. Given the possibility of trended variables, suggested in Table 5.11, two models were estimated for each country, including respectively an unrestricted [$H_0(r)$] and a restricted intercept [$H_1^*(r)$].

\textsuperscript{30} The ADF$_{SSC}$ test produced the same inferences, with zero lags in almost all regressions.
Table 5.12: Johansen cointegration tests - inflation $\pi_t^CPI$, tax rate $ln t_t$, base money $ln m_t$

<table>
<thead>
<tr>
<th></th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$H_1(r)$</td>
<td>$H_1^*(r)$</td>
<td>$H_1(r)$</td>
</tr>
<tr>
<td>order</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\lambda_{MAX}$ test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: r = 0</td>
<td>40.6481**</td>
<td>40.8622**</td>
<td>22.3955**</td>
</tr>
<tr>
<td>r ≤ 1</td>
<td>3.3088</td>
<td>3.8487</td>
<td>7.1640</td>
</tr>
<tr>
<td>r ≤ 2</td>
<td>-</td>
<td>-</td>
<td>2.5263</td>
</tr>
<tr>
<td>Trace test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: r = 0</td>
<td>43.9570**</td>
<td>44.7110**</td>
<td>32.0858**</td>
</tr>
<tr>
<td>r ≤ 1</td>
<td>-</td>
<td>-</td>
<td>9.6903</td>
</tr>
<tr>
<td>Cointegrating vector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ln t_t$</td>
<td>-0.2940 (0.0833)</td>
<td>-0.2953 (0.0834)</td>
<td>0.0665 (0.0460)</td>
</tr>
<tr>
<td>$ln m_t$</td>
<td>-0.5580 (0.1663)</td>
<td>-0.5580 (0.1663)</td>
<td>0.4017 (0.2038)</td>
</tr>
</tbody>
</table>

In the models for Italy, two point dummy variables were included (1951, 1974) to solve some problems of non-normality in one of the equations of the VECM. Standard errors in parenthesis. Critical values for 46 observations, $\lambda_{MAX}$/Trace tests (Pesaran, Shin and Smith, 1997) - $[H_1(r), 2 \text{ var.}]: 14.88/17.86 (5\%)$ and $12.98/15.75 (10\%)$ for $H_0: r=0$, 8.07 (5\%) and 6.50 (10\%) for both $H_0: r≤1$; $[H_1^*(r), 2 \text{ var.}]: 15.87/20.18 (5\%)$ and $13.81/17.88 (10\%)$ for $H_0: r=0$, 9.16 (5\%) and 7.53 (10\%) for both $H_0: r≤1$; $[H_1(r), 3 \text{ var.}]: 21.12/31.54 (5\%)$ and $19.02/28.78 (10\%)$ for $H_0: r=0$, 14.88/17.86 (5\%) and $12.98/15.75 (10\%)$ for $H_0: r≤1$, 8.07 (5\%) and 6.50 (10\%) for both $H_0: r≤2$; $[H_1^*(r), 3 \text{ var.}]: 22.04/34.87 (5\%)$ and $19.86/31.93 (10\%)$ for $H_0: r=0$, 15.87/20.18 (5\%) and $13.81/17.88 (10\%)$ for $H_0: r≤1$, 9.16 (5\%) and 7.53 (10\%) for both $H_0: r≤2$.

The hypothesis of no cointegration can be rejected, with more or less confidence, in all models considered above. However, in the cases of Germany and the Netherlands, the cointegrating vector indicates a negative coefficient of the tax rate variable (and also the others), contrary to the predictions of the model. Only in the case of Italy is this parameter positive, although not significant at the usual levels.\(^{31}\) These results suggest that only in Italy can inflation be considered an important element in the program of optimal taxation. In fact, within the countries included in the sample, Italy is the only country to have benefited from relatively higher seigniorage (see Table 5.1 above).

These results are very different from the ones obtained by Grilli (1989), using a similar testing methodology for a sample of ten EU member-countries during the period 1948-

\(^{31}\) The significance level increases if some extra dummy variables are added. To check the robustness of the results, the Engle-Granger two-step cointegration test was also performed, yielding similar results.
86. According to his results, the model is accepted for example in France, Germany and Italy, and rejected in the cases of Belgium, the Netherlands, the United Kingdom and others. This difference in the results obtained, which seems to arise from the longer time span used here, may reflect the alteration in the government’s priorities towards the exchange rate policy, in view of the EMU, at the cost of a ‘suboptimal’ inflation tax.

The possibility that the countries may have not followed the behaviour suggested by the theory because they have been committed to a near-fixed exchange rate regime which is their priority, is only one of the possible reasons for the almost complete rejection of the model. Other reasons may be advanced. For example, one assumption of the model is that the governments are constrained by the IBC, equation (5.19), a condition which, as seen in chapter 4, cannot be assumed a priori.

The model also assumes that the central bank produces base money at no cost and transfers it all to the government. However, Klein and Neumann (1990) show that there may be an important difference between the revenue produced by the central bank and the value transferred to the government (the size of this gap depends on institutional and operational factors). Goff and Toma (1993) show that this factor implies that the model of optimal taxation may not always involve a positive relationship between the inflation and the tax rates. Accordingly, Klein and Neumann also reject the conclusion from equation (5.23) that inflation is a positive function of real demand for base money relative to GDP.

In sum, the results of this simple model suggest that the theories of optimal taxation have not been generally applied in the EU. The monetary and fiscal authorities have not coordinated their policies with the aim of minimizing the distortionary costs of raising revenues. This may entail social costs, although the Commission of the EC (1990) has considered them to be minimal.
5.7. Summary of the main results

One of the main causes of concern with unsustainable fiscal policies is the conventional notion that increasing deficits and debt inevitably give rise to monetization and inflation. Although the empirical evidence is far from providing convincing support for this hypothesis, these concerns have recently been widely mentioned at the institutional level, as illustrated in the following passages,

"Many of us who are central bankers expect that a substantial reduction in the long-term prospective deficit of the United States will significantly lower very long-term prospective inflation expectations vis-a-vis other countries"

A. Greenspan, in Federal Reserve Bank of Kansas City (1995: p. 141)

"The failure to realise greater progress in budgetary consolidation [...] undermines price and exchange rate stability, increases uncertainty about the course of fiscal policy and erodes the credibility of policies. It contributes to an unbalanced policy mix, and it undermines the task of monetary policy"


The main objective of this chapter was to investigate whether this proposition of a positive effect of deficits on inflation could be ratified in the case of the EU during the post-war period. In fact, this was a central point in the discussion over the necessity of the Maastricht's fiscal convergence criteria to prevent excessive inflation in EMU.

The first sections of the chapter were reserved to the clarification, both conceptual and analytical, of some of the main expressions used in the text. The terms monetization, seigniorage and inflation-tax are often used interchangeably, albeit referring to related but quite distinct concepts. A clarification is also important in terms of the economic mechanisms presiding over the relation between the variables. The channels through which the government's budget interacts with the money creation and inflation are diverse, and the direction of causality is not always very clear. Throughout the chapter, several distinct effects have been identified: monetization (seigniorage), inflation-tax, wealth effect, cost of factors, real interest rate, and the tax/transfer system.

The empirical tests in this chapter were not able to find any evidence showing that the central government's deficits are monetized and ultimately lead to inflation, even though most of the countries analysed have been shown to present unsustainable fiscal policies.
On the contrary, the opposite relation was found in a few cases. In Belgium and Italy, seigniorage (measured by the ‘opportunity cost’ definition) affects positively the level of the deficit.

Furthermore, in most countries of the sample there is evidence of a unidirectional causal relationship running from the inflation rate to the budget deficit. This may be justified for example by the effects of inflation on certain categories of public revenues and expenditures. On the one hand, as noted by Persson et al. (1998), even if perfectly anticipated, inflation may reduce the deficit through the nominal features of the progressive tax/transfer system, by raising the real, effective, tax rates while eroding the real value of the transfers. On the other hand, it may also increase the deficit, if the Tanzi effect prevails.

One possible explanation for the lack of empirical evidence on the effects of deficits on monetization and inflation is the change of economic policy objectives since the early eighties, towards price stability and fiscal consolidation. In spite of the evident accomplishment in reducing inflation, the efforts on the fiscal side have not been so successful and, therefore, no statistically significant evidence was found of this relation.

On the contrary, this situation seems to have caused further problems. The decline of inflation may have not been perfectly credible, given the continuing fiscal problems. This inevitably causes an increase of the effective real interest rate on government debt (lower than expected inflation), an heavier debt burden, and therefore higher total deficits. This may explain why the results of the tests suggest a causal relationship running from the inflation rate to the deficit, not the contrary as suggested by the conventional economic theory. These results highlight the central role of the interest rate on the intertemporal budget dynamics, especially important in highly indebted countries such as Belgium and Italy. The next chapter examines in more detail this connection between the fiscal variables and the interest rate.

32 Fratianni and Spinelli (1999), for example, present evidence of this change of regime in the case of Italy. They report that the long period of ‘fiscal dominance’, in which the deficits determined the growth of the money supply, was reversed in the eighties and nineties, with the entry in the EMS, the higher independence of the central bank after 1981 and, more recently, the Maastricht Treaty. In fact, repeating the same empirical tests on the period 1950-1979 reveal a strong positive effect of the deficits on inflation, not found for the whole period (results not shown).

33 As noted by the European Commission (1995: p. 14), “prospects for the resolution of the fiscal imbalances remain uncertain, and fiscal convergence continues to be elusive”.

189
6. THE FISCAL EFFECTS ON THE INTEREST RATE

6.1. Introduction

The deterioration of the public finances in most developed countries during approximately the last two decades, raised concerns about its potentially negative effects on the economy. In particular, it prompted an extensive debate, largely focused on the United States, on the impact of the deficit/debt on inflation (examined in chapter 5) and the interest rate. This latter relationship has important consequences for the economy, in several complementary aspects. Directly, because of the effects on the 'real' macroeconomic variables such as investment, output and employment, as suggested by the standard macroeconomic theory. More indirectly, because of the central role of the interest rate on the stability of the budget dynamics.

If a government's expanding deficit or debt generates an increase in the interest rate, this would raise the debt servicing costs. Unless offset by a larger primary surplus or by monetization, these supplementary interest expenditures will further add to the total deficit and debt in a self-sustained, continuous process. The 'snowball effect' could rapidly transform a sustainable fiscal policy into an unsustainable one. This may be particularly relevant in the case of the Netherlands, for example, which, as seen before, does not satisfy the debt boundness condition (not strictly necessary) for sustainability.

In fact, the discussion over fiscal policy sustainability and its effects has arisen in parallel with the increase in the real interest rates during the 1980s. As can be observed in the analytical expression of the intertemporal budget constraint (IBC), the 'sustainable' levels of deficit and debt are directly dependent on the interest rate dynamics.

This chapter intends to examine the effects of a country's fiscal position on the level of nominal and real interest rates, interpreting the results by reference to the empirical observations of the previous chapters. The next two sections review some of the
economic theories which have been proposed to explain the connection between the fiscal variables and the interest rate, and survey the multitude of different empirical approaches to test it. Section 6.4 introduces the data employed in the empirical sections, and presents the methodology followed in the tests. Sections 6.5 and 6.6 exhibit the results of non-causality tests on the relationship between the interest rate and, respectively, the central government deficit and debt. The next section extends the tests to a broader level of government accounts, to investigate the influence of the local governments and the social security funds. Sections 6.8 and 6.9 examine whether the relationship is explained by domestic factors or primarily by European-wide fundamentals. The last section summarises the main findings and advances some economic policy implications.

6.2. Competing theories of the interest rate

The connection between deficits and interest rates has been explained by different economic theories, briefly reviewed in this section.\(^1\) The first group of theories, which has dominated the 'mainstream' view of the functioning of the economy, intends to explore the mechanisms through which the government deficits are positively transmitted to the interest rate. A second group in contrast, beginning with the Ricardian Equivalence theory, argues that the interest rates are determined independently of the fiscal policy actions.

6.2.1. The positive impact of deficits on the interest rate

The classical monetary theory of interest rates, first presented in 1898 by Wicksell,\(^2\) indicates that the level of the interest rate results from the arbitrage between the demand and supply of capital in the 'loanable funds market'. This idea can be formalised by the 'loanable funds model', associated with Robertson (1934). It models the interaction

\(^1\) For a more complete survey of the theory of interest rates see Green (1991).
\(^2\) First English translation in Wicksell (1936).
between the demand for capital or funds \((F)\), represented graphically by a negatively sloped curve \((D)\) reflecting the diminishing marginal productivity of capital, and the supply of capital, represented by a positively sloped curve \((S)\) reflecting the new savings available in the market (Figure 6.1).

The intersection point of the demand and supply schedules indicates what Wicksell denominated the ‘natural interest rate’, since it is supposed to be determined by real forces of productivity and savings, and to equal the rate of rentability of the productive capital. It is “the same as the rate of interest which would be determined by supply and demand if no use were made of money” (Wicksell, 1936: p. 102).

The natural rate and the market real interest rate are not necessarily equal, at least in the short-term, namely due to the central bank’s control over the money supply. However, when the two rates differ, an equalisation mechanism is immediately set in motion. For example, starting from an equilibrium situation \((ir_1, F_1)\), suppose that a fall in the marginal productivity of capital causes a downward shift in the demand schedule and the consequent decline of the natural rate. If, for any reason, the market interest rate does not accompany this fall, there will be an excess supply of funds \((F_1-F_2)\), certain investment projects will have to be abandoned, and income and employment will fall. This causes an upward shift in the supply schedule until the natural and the market interest rates meet again and the equilibrium is restored.
According to this theoretical framework, an increase in the government's total deficit, implying necessarily additional borrowing in the market, will shift the demand curve upwards. To restore the equilibrium in the market, the natural rate of interest must rise. This increase in the interest rates may however be avoided if the private demand for funds declines sufficiently to offset the increase in public demand, if the supply curve moves downward (e.g. if the Ricardian equivalence theorem holds), or if the supply curve is infinitely elastic (e.g. in a small economy with open capital markets). These 'special cases' will be presented in the next section.

The classical view was partly challenged by Keynes' (1936) 'liquidity preference theory', usually examined within the IS/LM analytical framework. According to this theory, the interest rate is the equilibrium price between the supply of money, or liquidity, supposedly controlled by the monetary authorities, and the demand for money, a function of income and the nominal interest rate. The demand for money depends on the various reasons to detain liquidity, including the so-called 'speculative motive', which is conditional on the demand for alternative financial assets.

Like the classics, Keynes assumes a fixed exogenous quantity of money. The fundamental difference between the Classical and the Keynesian theories of interest rate determination is that in the former the interest rate is essentially a 'real' phenomenon while in the latter it is a monetary phenomenon, determined in the money market and then transmitted to the real part of the economy. Causality runs in opposite directions in both theories.

In sum, according to the Keynesian theory, an expansionary fiscal policy, for example, implies an increase in the demand for money and output, requiring a rise in the interest rate to restore equilibrium in the markets (unless the economy is caught in the 'liquidity trap', in which case the nominal interest rate is so low that the opportunity cost of holding money is negligible).

The more recent Neoclassical perspective, which may be examined within any of the two theoretical constructions presented above ('neoclassical synthesis'), focuses particularly on the economic 'crowding-out' consequences of deficits. According to the standard view, an increase in the deficit, with the consequent additional government borrowing, reduces the total amount of national savings available, requiring a higher interest rate to
restore the equilibrium. The interest increase will change the final composition of output, depending on the degree of economic openness and international capital mobility. In a closed economy, for example, it will crowd-out private investment, since the government competes with private individuals for the available savings.

In a small open economy, with high capital mobility, it will crowd-out net exports. The interest rate increase will attract a capital inflow, causing the domestic currency to appreciate and consequently leading to a trade balance deterioration. At the same time, this inflow of capital forces a fall in the interest rate, limiting to a certain extent the crowding-out effect on domestic private capital formation. In both cases, whatever the economic characteristics of the country, the expected final result is inevitably slower economic growth.

Finally, another alternative approach claims that the positive effect of the fiscal variables on the interest rate depends, at least in part, on the efficient functioning of the financial markets. The ‘market discipline hypothesis’ maintains that high levels of deficits and especially of debt induce the financial markets to include (or increase) a default risk premium on the interest rate demanded on government bonds. Some authors even claim that the financial market is the first place to look for a ‘natural’ restraint on fiscal indiscipline (Bishop et al., 1989, Bishop, 1990, and Fitoussi et al., 1993, among others).³

Although this theory may help explain the interest rate differentials within a monetary union such as the US or the EMU,⁴ it does not seem able to explain the relation between the fiscal variables and the interest rate in the whole sample considered in this study. The functioning of market discipline requires the verification of a group of conditions (Lane, 1993), such as the perfect integration of capital markets and the non-monetization of the debt, hardly met in the whole period analysed here.

The empirical evidence on this theory, almost exclusively centred on the United States, provides mixed results. While Goldstein and Woglom (1992) and Bayourmi et al. (1995)
present results supporting the theory, Eichengreen (1990) and Alesina et al. (1992) claim that the risk premium is too small to matter.

6.2.2. Contrasting views: no fiscal effects on the interest rate

Contrasting markedly with all the above theories, the ‘Ricardian Equivalence Theorem’ (RET) on intertemporal government finance claims that no relationship is to be expected between the government deficits and the interest rate. According to this theory, the rational private individuals interpret the present government deficits as an intertemporal reallocation of taxes. They fully anticipate the effects of the current deficits and consequent debt accumulation on their future tax liabilities, and therefore increase their savings now in an amount exactly equivalent to the present value of these future tax liabilities. Their optimal consumption plans will not be altered.

In consequence of this behaviour the government deficit, which implies public dissaving, will be completely offset by a compensating increase of private savings, neutralising the impact of the deficit on the interest rate. Coincidentally, Belgium, Italy and the Netherlands, the three most indebted countries in the sample, present average private savings relative to GDP, in the period 1960-96, above the European Union’s average and well above any of the three other countries in the sample.6

Many years after its first enunciation and immediate dismissal by Ricardo, several economists have returned to this hypothesis. Tobin (1952: p. 117) asks whether “the additional taxes which are necessary to carry the interest charges reduce the value of other components of private wealth?”. Bailey (1962: p. 75) admits the possibility that “households regard deficit financing as equivalent to taxation”. He later enumerates the conditions necessary for a null net wealth effect: “If all households accurately foresee their future disposable incomes, including the tax implications of current fiscal policy, and if they all plan gifts and bequests to their heirs, taxation and debt finance are equivalent” (Bailey, 1972: p. 651).

---

5 This expression was first proposed by Buchanan (1976).
6 Respectively 22.6%, 27.4% and 23.5% of GDP, while the EU(12) average for the same period was 21.2%. The data is from European Commission (1998) and refers to the “Gross saving of the private sector as a percentage of GDP at market prices”.

195
But it is Robert Barro who demonstrates this proposition analytically, within the theoretical framework of an overlapping-generations model. Barro (1974: p. 1116) concludes that "fiscal effects involving changes in the relative amounts of tax and debt finance for a given amount of public expenditure would have no effect on aggregate demand, interest rates, and capital formation". The RET is a sufficient, although not necessary, condition to explain the lack of relationship between deficits and interest rates.\(^7\)

The main objections to this theory focus on the large set of strong assumptions on which it is theoretically grounded. As underlined above by Bailey, it is necessary that the individuals are rational, have perfect foresight, are capable of understanding completely the intertemporal budget constraint they face, and also the existence of an operative, altruist, intergenerational transfer mechanism. Furthermore, it is also necessary to assume perfect capital markets and that taxes are lump-sum and nondistortionary.\(^8\)

In spite of all these restrictions and seemingly unrealistic assumptions, some authors followed Evans' (1985: p. 85) advice to "judge the utility of an assumption primarily by its predictive and explanatory power and not by its realism". In fact, the RET was supported by several empirical studies during the 1970s and 1980s, all for the United States, which have found no evidence that private consumption reacted to a deficit-caused increase in disposable income (see, for example, Kochin, 1974, Kormendi, 1983, Aschauer, 1985, and Carroll and Summers, 1987).\(^9\)

This idea can be easily explained in a very simple model of consumption optimisation (the necessary simplicity of the model insinuates how strong the conditions of the RET must be). Suppose that individuals maximise their lifetime (to simplify, here divided in two periods) utility from consumption subject to the following income restrictions, in present value,

\[
C_1 + C_2 \delta = Y_1 - T_1 + (Y_2 - T_2) \delta ,
\]

\(^7\)Conversely, the hypothesis of a positive relation between deficits and interest rate has been used as a way of testing the Ricardian Equivalence Theorem. However, as shown by Rose and Hakes (1995) and Nunes and Stemisiotis (1995), rejection of this hypothesis is only a necessary but not a sufficient condition for Ricardian equivalence.

\(^8\)For a discussion of these assumptions see, for example, Barro (1989) and Seater (1993).

\(^9\)Contrary evidence, also for the US, is reported for example by Feldstein (1982).
where $C_i$, $Y_i$, $T_i$ represent, respectively, consumption, income and taxes in period $i$ ($i=1,2$), and $\delta$ is the discount factor. The RET assumes that individuals incorporate the government's intertemporal budget constraint (it is assumed that the government repays all debt, the initial amount $D_0$ plus the debt meanwhile incurred in period 1 - no Ponzi games)

\begin{equation}
D_0\delta^{-1} = \sum_{j=1}^{2} \delta^{j-1}(T_j - G_j)
\end{equation}

in their consumption optimisation problem (as before, $G$ represents the government expenditures). Substituting (6.2) in (6.1) yields

\begin{equation}
C_1 + C_2\delta = Y_1 + Y_2\delta - (G_1 + G_2\delta).
\end{equation}

Equation (6.3) shows that the government expenditures are considered in the private individual's consumption decisions, but the way they are financed, through debt or taxes, is irrelevant. Individuals perceive deficits as postponed taxes, not affecting their intertemporal consumption plans.

In the 1990s there was the possibility of a 'Maastricht effect' in the European Union, somehow similar in its outcome to the Ricardian Equivalence Theorem. An individual who observes an unfavourable fiscal situation in his country, knowing that the Maastricht's fiscal criteria will have to be met in the medium-term, expects a tax increase in the near future, and may increase savings accordingly to account for that tax increase. The final effect is similar to the RET, with the increase in private savings compensating at least part of the government dissaving with no effects on the interest rate. This case also relaxes some of the strong assumptions of the Theorem, since it requires a smaller time interval before the taxes are expected to be collected.

More recently, another theory has been put forward to explain the possible lack of relationship between government deficits and interest rates, when foreign demand is

---

10 Some authors (Barro, 1987, Barth and Bradley, 1989, Swami et al., 1990, Dua, 1993, Vamvoukas, 1997, among others) include the government expenditures, together with the deficit, in the model used to test empirically the influence of fiscal policy on the interest rates. However, as pointed out by Seater (1993: p. 176), this procedure of including government expenditures in the regression of the interest rate (and excluding revenues) raises problems of omitted variable bias and may indicate negative relations between these variables.
considered. According to some authors (Tran and Sawhney, 1988, Monadjemi and Kearney, 1991, Nunes and Stemitsiotis, 1995, Knot, 1996, and Tanzi and Fanizza, 1996 among others), this may be due to the growing integration of capital markets, widening the area where governments can obtain financing, and therefore easing the pressure on domestic interest rates. This is particularly conceivable in a small open economy, where an increase in the government's borrowing needs could be met by foreign capital inflows at the world interest rate, without affecting the domestic rate.

This 'capital inflow hypothesis', suggested by Darby (1979) in the similar context of social security financing, has been widely discussed and tested in the literature, since the often quoted paper by Feldstein and Horioka (1980). They revealed the existence of a strong empirical correlation between domestic investment and savings, contrary to the common widespread belief of international financial integration. In spite of the large controversy inspired by this paper, and the huge amount of empirical research in this area, no definitive conclusions have yet been reached (see, for example, Lemmen and Eijffinger, 1995).\textsuperscript{11}

Together with the RET, these theories illustrate two of the above mentioned exceptions to the functioning of the classical loanable funds model illustrated in Figure 6.1. While the RET implies a downward shift in the supply schedule, due to the increase in private savings, the integration of the capital markets suggest a supply curve infinitely elastic. In both cases, the upward shift in the demand schedule due for example to a government deficit will not necessarily cause an interest rate increase.

This section has, so far, described several theories suggesting a positive impact of the deficits on the interest rate, followed by other views justifying the non-existence of such relation. In contrast to all these approaches, some post-Keynesian authors call the attention for the predominance of the opposite causal relationship, running from interest rates to deficits (see, among others, Horstman and Schneider, 1994, Smithin, 1994, and Wray, 1997).

\textsuperscript{11} For the particular case of the EU, Pentecost and Holmes (1995) and Holmes and Pentecost (1996, 1997 and 1999), for example, present evidence of an increasing degree of financial integration, particularly in the countries more committed to the ERM.
As implied above in the loanable funds model, and also in the original IS/LM framework (see for example Keynes, 1979: p. 222), the relative availability of credit in the market plays a fundamental role in the determination of the interest rate. According to this model, there is an exogenous, fixed, supply of savings and money in the economy, and consequently the level of investment depends on the total pool of savings available at a certain time. The interest rate is an endogenous variable, resulting from the interaction between the demand and supply of savings (or money, for the keynesians) in the market. This view is completely reversed by the post-keynesians, for whom the interest rate is an exogenous variable, determined by the decisions of the central bank and therefore not influenced by the government’s deficit. The money supply is the endogenous variable, out of the control of the central bank.

The theoretical framework for this hypothesis is provided by the ‘horizontalist’ approach, as advocated by Kaldor (1982) and Moore (1988), and by the Franco-Italian ‘circuit school’ of Graziani (1990) and LeBourva (1992). In a credit economy, the money supply is a demand-driven variable, consisting basically of commercial bank’s deposits. It varies endogenously with the volume of borrowing requested from the banks, which implies that the investment decisions can precede savings availability. The central bank, being the lender-of-last-resort, provides the reserves necessary to accommodate the change in the demand for credit, although at the interest rate it chooses.\textsuperscript{12}

In sum, according to this theory, the interest rates are set by the central bank’s policy decisions and not by the government’s fiscal actions. On the contrary, a tight monetary policy for example, leading to an increase in the interest rate, is positively transmitted to the government deficit through two main channels. Directly, and depending on the level of the debt, through the increase in interest expenditures. Indirectly, the higher interest rate slows down the economy and, through the ‘automatic stabilisers’, reduces tax revenues and increases certain transfers (social transfers, such as unemployment benefits, and possibly also for those sectors of the economy more sensitive to interest rate increases).

\textsuperscript{12} This assumption about the functioning of the central bank is confirmed, for example, by Goodhart (1989: p. 208) and Goodfriend (1993: p. 3), both directly involved in the central bank’s operational procedures.
6.3. Review of the empirical literature

Given the variety of opposite views suggested by the different economic theories surveyed above, this question becomes ultimately an empirical issue. However, in the empirical literature this relation remains highly controversial. Numerous tests on the existence of a positive relationship have been performed since the mid 1980s, but the conclusions achieved are almost perfectly divided into a rejection or acceptance of the hypothesis. Extensive lists of papers could be presented, showing opposite results, but it is difficult to derive meaningful and useful comparisons. On the one hand, most empirical literature in this area has been confined to the United States, which may not necessarily be indicative of the experiences of the countries here analysed. On the other hand, even for the same country, meaningful comparisons are difficult for the well-known reasons of differences in the sample period, econometric estimation methods, model specification, variables involved, among many others.

The variables included in the estimating equation to explain the interest rate, apart from the deficit and/or debt, vary widely according to the authors and the competing economic theory adopted. Some authors simply choose arbitrarily some variables to explain the interest rate movements, and estimate a single-equation regression (Craig, 1991 and 1994 or Laumas, 1989, for example), or a VAR model (Arora and Dua, 1995, Miller and Russek, 1996, Vamvoukas, 1997 and Cheng, 1998, for example). Most tests, however, in the tradition of the 1960s Cowles Foundation structural approach, begin by adopting as reference one of the theoretical models presented above, and then estimate a non-dynamic reduced-form single equation of the interest rate, examining the sign and significance of the coefficient of the deficit or debt variables.

\textsuperscript{13} Other authors prefer to rely on more informal, graphical or descriptive, analysis. See, for example, Friedman (1992), Tanzi and Fanizza (1996), Christiansen and Pigott (1997) or Caselli, Giovannini and Lane (1998).

\textsuperscript{14} The loanable funds framework, for example, is used to describe the interest rate determination process by, among others, Hoelscher (1986), Mehra (1992), Nunes and Stermitzis (1995), Gupta and Moazzami (1996) and Cebula (1997a,b,c, 1998). Other authors prefer an IS/LM framework: Evans (1985, 1987), Al-Saji (1993), Knot (1996) and Allen and Wohar (1996), among many others. An alternative which has been used as a theoretical background model to derive a reduced-form equation to be estimated, is the 'overlapping generations model', which pretends to synthesise the previous main economic schools and incorporate the 'mainstream' macroeconomic thought (see, for example, Blanchard and Fischer, 1989). While the above are structural or equilibrium models, this one is an optimization model, based on the
Following the analytical expressions from the IS/LM model for example, or even from the simple optimization model of the previous section, some authors have claimed that the results of the empirical tests may be distorted by the influence of monetary policy actions and international capital flows. The central bank's interventions to affect the interest rates, and the international movements of capital may offset the interest rate responses to the government's deficits and debt.

The proposed alternative is to test the connection between the interest rate and the fiscal variables indirectly, through the effects of the latter on a third variable, closely related to the interest rate. Bahmani-Oskooee and Payesteh (1994), for example, analyse the impact of deficits on capital flows, while Evans (1985), Gulley (1994) and Kalulumia and Lukusa (1996) examine the effects of deficits and/or debt on money demand (all these studies use US data). However, like in all other approaches, the results are not unanimous. While Evans and Gulley were not able to find any (indirect) relationship between the fiscal variables and the interest rates, the other authors mentioned above report a positive link between them.

One potential econometric problem with most of these empirical studies is the non consideration of the time series properties of the underlying series, resulting potentially in wrong specifications of the models, and spurious results (among the most recent studies, see, for example, Al-Saji, 1993, Nunes and Stemitsiotis, 1995, Belton and Cebula, 1995, Gupta and Moazzami, 1996, and Cebula, 1997b). Other studies acknowledge this fact and first-difference all I(1) variables prior to estimation, loosing potentially important short-term information (this is the case of, for example, Allen and Wohar, 1996 and Cebula, 1998).

One other weakness of most empirical approaches to this question is that they search for correlation without paying attention to issues of causality. As can be seen in Figure A.10 in the appendix, there seems to be a positive correlation between the nominal interest rate and the total deficits and debt in some countries (this empirical link is much less evident with the real interest rate). However, the existence of correlation does not necessarily imply that the effect runs exclusively from the fiscal variables to the interest rate or even maximisation of the household's lifetime utility. Among others, it was used by Pittaluga and Vaciago (1994), Tanzi and Fanizza (1996) and Mongelli (1997).
that they are causally linked (Granger, 1988). It is necessary to take into account the possibility of an inverse causal effect. The level of the interest rate has an important influence on the deficit, mainly through the interest expenditures, and consequently on the accumulation of public debt.

Only a few empirical studies have specifically addressed the question of causality in the relation between the fiscal variables and the interest rate. Using the more standard causality tests, for example, McMillin (1986), Darrat (1989 and 1990) and Miller and Russel (1996) found no evidence that deficits cause interest rates. In fact, the first two found evidence of the opposite causal effect, that the interest rate Granger-causes the deficit. However, using the more appropriate VECM procedure, Miller and Russel (1991) and Arora and Dua (1995) claim that the government deficits cause the interest rate. All these studies refer exclusively to US data.

6.4. Data and Methodology

This section is a complement to chapter 3, describing the variables and the methodology specifically employed in this particular chapter. The first part justifies the choice of variables, examines their evolution and performs some tests to assess their statistical properties. The second part clarifies the approach followed in the empirical tests, and establishes the main hypothesis being investigated.

6.4.1. Choice of variables and other data issues

This chapter examines the relationship between the government’s total deficit or debt and the interest rate, both in nominal and real terms.\(^{15}\) There is no consensus on which is the most appropriate measure of the variables used in this research area, and therefore one

---

\(^{15}\) Although most studies choose to analyse the effects of the fiscal variables exclusively on either the nominal or the real interest rate, Allen (1991), for example, has shown that the specification of the dependent variable has non-negligible effects on the results of the tests. He finds that deficits and debt do not have a positive effect on the nominal interest rate, but have on the real rate.
possible procedure is to use alternative concepts and check whether the results are
sensitive to the different measures. However, given the number of countries considered,
and the time span covered, the alternatives are relatively few if one aims at maintaining
the results comparable. The question of which is the most adequate measure of the deficit
and debt was already discussed in chapter 3, and therefore this section will be concerned
primarily with the interest rate variables.

The first question to address is what type of nominal interest rate should be chosen, in
terms of maturity, for example. Although most empirical studies have used short-term
rates, several arguments may be put forward in favour of the use of a long-term rate. The
long-term interest rate has an important role in economic intertemporal decision making,
particularly those related to investment. Macroeconomic theory suggests that the effects
of the deficits on the economy are transmitted through the long-term interest rate, not the
short-term rate, because the interest sensitive part of the private investment (such as
housing, factories or industrial equipment) depends on the long-term interest rate.
Besides, movements in the short-term rates are more volatile and difficult to interpret,
since they are highly influenced by transitory factors and by the monetary authorities

For the sample period and the countries considered, there seems to be only two
alternative measures of the nominal interest rate available. One is the most commonly
used long-term government bond yield. The other is the implicit average interest rate paid
on the government debt, computed as the ratio of total interest expenditures to the total
stock of debt, usually designated as ‘implicit rate’.17

The implicit interest rate is usually below the market rate (see, for instance, OECD,
1990), which may be explained by the existence of non-interest bearing (or below the
market rate) debt acquired by the central bank.18 The difference between the two series
depends also on the average maturity of the government debt, since longer-term debt may
have been issued in the past at interest rates distinct from the current ones. As a result, the

16 Anyway, Passet (1997) and Christiansen and Pigott (1997) argue that the periods where the long-term
rates diverge from the short-term ones are ‘very rare and ephemeral’.
17 Chosen for example by Uctum and Wickens (1996), Caselli, Giovannini and Lane (1998), and Artis and
18 An alternative explanation is that the interest payments are underestimated because they include only
coupon payments, disregarding for example the zero coupon debt (Uctum and Wickens, 1996).
implicit interest rate may not be representative of the present market conditions, distorting the conclusions of the tests. Besides, there is no complete and reliable series of interest expenditures for the time span considered here. Therefore, this chapter will use only the long-term interest rate on government bonds.

An alternative, or complement, to using the level of each country’s interest rate, is to substitute it by the interest rate differential in relation to some basis country (usually Germany, the United States, or the average of the group), and study the relationship in relative terms. This procedure will be applied in section 6.8.

Another important choice to be made is how to estimate the real interest rate. The theoretically most correct procedure is to use the expected real interest rate. The problem is that these expectations are not observable in practice, and the usual alternative procedure is to employ some proxy for expectations (forecasts from VAR’s, trend inflation using the Hodrick-Prescott filter, rolling moving-averages of past values, indexed bonds, and several others). However, there are always some reservations about the quality of these proxies and its effect on the final results of the tests. Furthermore, the main problem in this particular area of research is that “measurement errors in the proxies for expectations biases the estimated coefficients toward zero, and toward the null hypothesis of Ricardian Equivalence” (Elmendorf and Mankiw, 1998: p. 47). This chapter will therefore use the \textit{ex post} real interest rate, computed as the difference between the nominal long-term rate and some \textit{ex post} measure of the inflation rate. Assuming rational expectations, this rate is an unbiased estimate of the \textit{ex ante} real interest rate (Huizinga and Mishkin, 1984).

Two alternative price indices were considered. The initial approach was to use the GDP deflator, obtained from the series of GDP at current and constant prices. However, the latter series was not consistently defined over the entire sample, and a different data source was used to fill the gaps, demanding a careful sensitivity analysis of the results. Besides, these series present some breaks due to methodological adjustments in the way they are estimated in each country, which have taken place several times during the sample period considered. All these factors have a non-negligible effect on the results of

---

\textsuperscript{19} This variable would have to be taken from the OECD until 1968 and the IMF since 1969, and then divided by the total debt series of the IMF.
the tests (not presented) and therefore the GDP deflator series was finally discarded. Alternatively, the inflation rate was measured as the growth rate of the consumer price index.

A third choice to be made concerns the time frequency of the observations. Following the usual approach of the previous chapters, mainly motivated by the problem of stochastic seasonality highlighted in section 4.6.2, the immediate tendency is to continue using annual data. This choice is supported, in this particular chapter, by the suggestions that higher frequencies bias the results towards finding no significant relationship between deficits and interest rates (as pointed out by Hoelscher, 1986). Apart from the problems of seasonality already mentioned, this phenomenon may be justified by the existence of portfolio adjustment lags, or because next month's or quarter's deficit (more easily predicted than with annual data) is already influencing the present value of the interest rate. Annual data is less prone to be distorted by temporary shocks and therefore may allow a better analysis of the fundamental factors affecting the interest rate.

Finally, all the tests will be performed on two alternative levels of government accounts, the central and the general government. A comparison between them reveals for example the importance of the social security funds in the analysis. As observed in chapter 4, these funds play a non-negligible role in questions of fiscal policy sustainability in some countries, and are regarded by many authors as the main source of fiscal problems in the medium-term for the developed countries.

Table 6.1 presents unit root tests on all the variables used below in the empirical sections, except for the total surplus of the central and general government, already presented in sections 4.2 and 4.6.1, respectively. The ADF_{SBC} and the ADF_{ST} tests are complemented by the test suggested by Perron and Vogelsang (1992a) and extended in Perron (1994), designated by PV, which allows for the presence of a structural break in the level of the series, in a date endogenously chosen. The hypothesis of a unit root on all the first-differenced series was also tested (results not shown) and clearly rejected in all cases.

The different types of test present in general coincident results. The very few cases where unanimity does not happen will be carefully examined in the empirical applications of the next sections. As shown in Table 6.1, the nominal interest rate series (and differentials)
present in general a nonstationary behaviour, while the real interest rate series are, with a few exceptions, stationary.

### Table 6.1: Unit root tests on the series of debt and interest rates

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>U. Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>nominal interest rate</strong> $(i_t)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ADF_{SBC}$</td>
<td>-1.9455 (1)</td>
<td>-1.8120 (1)</td>
<td>-3.9241  (1)</td>
<td>-1.9858 (1)</td>
<td>-1.7894 (0)</td>
<td>-1.7918 (1)</td>
</tr>
<tr>
<td>$ADF_{ST}$</td>
<td>-1.9455 (1)</td>
<td>-1.8120 (1)</td>
<td>-3.9241  (1)</td>
<td>-1.9858 (1)</td>
<td>-1.9104 (1)</td>
<td>-1.7918 (1)</td>
</tr>
<tr>
<td>PV</td>
<td>-2.6257 (1)</td>
<td>-2.8157 (1)</td>
<td>-4.8756  (1)</td>
<td>-3.5011 (1)</td>
<td>-3.6528 (1)</td>
<td>-2.6589 (1)</td>
</tr>
<tr>
<td>[1973] [1972] [1969]</td>
<td>[1974]</td>
<td>[1965]</td>
<td>[1968]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>real interest rate</strong> $(r_{t, CP}^i)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ADF_{SBC}$</td>
<td>-3.5953  (0)</td>
<td>-3.9163  (0)</td>
<td>-2.8377  (2)</td>
<td>-3.3588  (0)</td>
<td>-1.1490 (2)</td>
<td>-3.3107 (0)</td>
</tr>
<tr>
<td>$ADF_{ST}$</td>
<td>-3.5953  (0)</td>
<td>-3.9163  (0)</td>
<td>-2.8377  (2)</td>
<td>-3.3588  (0)</td>
<td>-1.1490 (2)</td>
<td>-3.3107 (0)</td>
</tr>
<tr>
<td>PV</td>
<td>-5.6063  (3)</td>
<td>-6.3197  (2)</td>
<td>-2.9262  (2)</td>
<td>-3.1103 (3)</td>
<td>-3.6517 (2)</td>
<td>-4.8036  (0)</td>
</tr>
<tr>
<td>[1977] [1981] [1982]</td>
<td>[1981]</td>
<td>[1977]</td>
<td>[1981]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>total debt, central government</strong> $(d_t)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ADF_{SBC}$</td>
<td>-1.4201  (1)t</td>
<td>0.8181   (0)t</td>
<td>-1.4119  (0)t</td>
<td>-0.8106 (0)t</td>
<td>-1.1841 (3)</td>
<td>-1.1775  (1)t</td>
</tr>
<tr>
<td>$ADF_{ST}$</td>
<td>-1.4201  (1)t</td>
<td>1.0021   (1)t</td>
<td>-1.4119  (0)t</td>
<td>-0.8106 (0)t</td>
<td>-1.1841 (3)</td>
<td>-1.1775  (1)t</td>
</tr>
<tr>
<td>PV</td>
<td>-2.8453  (1)</td>
<td>-3.5785  (1)</td>
<td>-2.6729  (1)</td>
<td>-2.9321 (1)</td>
<td>-1.5940 (1)</td>
<td>-2.5405  (1)</td>
</tr>
<tr>
<td>[1975] [1975] [1970]</td>
<td>[1971]</td>
<td>[1965]</td>
<td>[1970]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>nominal interest rate differential vis-à-vis US/Germany</strong> $(i_d^D)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ADF_{SBC}$</td>
<td>-2.0210  (0)</td>
<td>-1.9448  (0)</td>
<td>-1.9202  (0)</td>
<td>-2.4484 (1)</td>
<td>-3.5907  (1)</td>
<td>-1.7799  (0)</td>
</tr>
<tr>
<td>$ADF_{ST}$</td>
<td>-2.0210  (0)</td>
<td>-1.9448  (0)</td>
<td>-1.9202  (0)</td>
<td>-2.4484 (1)</td>
<td>-3.5907  (1)</td>
<td>-2.1809  (1)</td>
</tr>
<tr>
<td>PV</td>
<td>-2.7490  (1)</td>
<td>-2.5605  (0)</td>
<td>-3.4576  (0)</td>
<td>-4.1671 (1)</td>
<td>-4.6661  (2)</td>
<td>-3.2190  (1)</td>
</tr>
<tr>
<td>[1976] [1976] [1974]</td>
<td>[1975]</td>
<td>[1983]</td>
<td>[1973]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>real interest rate differential vis-à-vis US/Germany</strong> $(r_d^D)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ADF_{SBC}$</td>
<td>-2.7556  (0)</td>
<td>-4.8955  (0)</td>
<td>-4.5573  (0)</td>
<td>-2.5873 (0)</td>
<td>-5.7084  (0)</td>
<td>-3.6119  (0)</td>
</tr>
<tr>
<td>$ADF_{ST}$</td>
<td>-3.8255  (3)</td>
<td>-1.9606  (3)</td>
<td>-2.9145  (2)</td>
<td>-2.5873 (0)</td>
<td>-5.7084  (0)</td>
<td>-3.6119  (0)</td>
</tr>
<tr>
<td>PV</td>
<td>-5.2451  (3)</td>
<td>-5.5180  (0)</td>
<td>-4.8547  (1)</td>
<td>-3.0298 (3)</td>
<td>-6.1340  (1)</td>
<td>-4.8937  (0)</td>
</tr>
<tr>
<td>[1961] [1984] [1980]</td>
<td>[1971]</td>
<td>[1977]</td>
<td>[1970]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The number of lags (in parenthesis) is chosen by the Schwarz Bayesian Criterion ($ADF_{SBC}$) or starting from a maximum lag length (3) and removing the last lagged term if not significant at the 10% level ($ADF_{ST}$ and PV) - in all cases, ensuring no serial correlation or heteroscedasticity in the residuals. All models include a trend, if significant. The PV structural break test is based on an AO model with a break in the level or the trend (for $d_t$), the test statistic is chosen to maximise $I_{d}$ in equation (4.3), the date (in square brackets) is chosen to maximise $I_{d}$ in (4.2). The asterisks (*), (**), and (***) indicate that the null hypothesis of a unit root was rejected at the, respectively, 10%, 5% and 1% level of significance.

The presence of unit roots in real interest rates has important macroeconomic implications. The standard economic theory maintains that in the long-run, equilibrium
real interest rates should equal the marginal productivity of capital with the marginal rate of time preference of the individuals, i.e., equal savings and investment. This equilibrium long-run real interest rate is expected to be stationary. However, the empirical research, largely focused on the United States case and the testing of the Fisher relation, provides mixed results. While Rose (1988) concludes that the real interest rate must be I(1), Mishkin (1992), Crowder and Hoffman (1996) and Lai (1997) suggest the contrary. Perron (1990) finds that the US real interest rate becomes stationary only after allowing the existence of a structural break in the series. This possibility was considered in Table 6.1, without significant changes in the conclusions. In the case of the European Union, Germany seems to be a notable exception in terms of stability of the nominal interest rate, as also reported by Gupta and Moazzami (1996) for the period 1967-90.

As can be detected in Figure A.10 in the appendix, some of the series present a rather unusual volatile behaviour in the beginning of the sample. Consequently, all tests below were repeated in a truncated series to assess the sensitivity of the results and, when necessary, an intervention dummy was included in the model to account for these outliers.

The plots also suggest that most variables may present a structural change in the beginning of the 1970s. To check the sensitivity of the results to these mid-sample changes, three alternative strategies were implemented. The first approach consists merely of including dummy variables (if statistically significant) in the model, to account for the changes in the level/trend of the variables, in the dates suggested by the Perron and Vogelsang’s (1992a) tests presented in Table 6.1, or by a visual inspection of the residuals.

The second approach results from the observation that the structural changes in the series roughly coincide with the oil shocks of the 1970s. Therefore, and following the suggestion of Johansen and Juselius (1992: p. 219), an exogenous and stationary variable, the changes in the oil price, is included in the deterministic part of the model.20 The variable used in the tests is the price of Saudi Arabia’s crude petroleum in dollars,

20 More precisely, they include in the model the variables \( \Delta po_t \) and \( \Delta po_{t-1} \), where \( po \) stands for the world price of crude oil.
multiplied by the exchange rate to transform into domestic currency, and divided by the domestic price index.

The third approach consists of repeating the tests on a sub-sample restricted to the period 1972-1996. The choice of this particular period is motivated by the major macroeconomic events coinciding with the beginning of the sub-sample and affecting all the countries considered: the end of the Bretton Woods system in 1971;21 the beginning of the ‘Snake Arrangement’ in 1972; the EU enlargement to nine members in 1973; the energy crisis of 1973. It also coincides with a major change in the trend followed by the stock of debt. Since 1950, the huge financial burden left by the second World War in most countries had been reducing slowly. In the early 1970s this trend was reversed in all countries, as can be observed in Figure A.10, for example.

This third procedure has the advantage of avoiding the outliers and structural breaks on the beginning of the sample and the beginning of the 1970s. It also allows, by comparison with the results for the full sample, to observe whether the relationship between the fiscal variables and the interest rates has changed in the more recent period. However, given the small number of observations in this restricted sample, the results of the tests must be carefully interpreted.

In order to avoid an unreasonable number of tables, only the main test results will be presented in this chapter. The innumerable supplementary tests intended to check the sensitivity of the main results will only be referred to when significantly contradicting the main conclusions.22

6.4.2. Methodology adopted to test the relationship

The empirical studies of the relationship between the fiscal variables and the interest rates, discussed in section 6.3 above, have been in general based on structural, large-scale, macro-econometric models or, more frequently, on reduced-form equations. The

21 Heinemann (1998) investigates the correlation between the fiscal variables and the alteration in exchange rate system, and its possible implications for the present major transition of exchange rate system in Europe, with EMU.
22 One of the motivations for all these additional tests is that the empirical results shown below go somewhat against what would be expected, according to the standard macroeconomic theory, and must therefore be examined with extra caution.
former can be immediately excluded from the present study on the grounds of data availability and small number of degrees of freedom for the sample period and data frequency analysed here. Besides, it would be unreasonable to use a large-scale structural model to examine the mere partial relationship between two variables. On the other hand, the reduced-form models have the important disadvantage, among others, of assuming exogeneity of all dependent variables to the interest rate. As noted above, the level of the interest rate may influence the level of the deficit, through its direct effects on the interest expenditure or its indirect effects on the primary deficit.

There is therefore an identification problem in most empirical work, since it may be difficult to separate the effects of the deficits from their causes. In fact, at least four distinct outcomes may arise from the empirical analysis of the relationship between deficits and interest rates, with different economic consequences and policy implications (Table 6.2).

**Table 6.2: Possible outcomes of the tests**

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Theoretical background</th>
<th>Main consequences</th>
<th>Policy implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>I def → ir</td>
<td>Classical, monetarist view</td>
<td>crowding-out of private investment and net exports, slower economic growth</td>
<td>need for fiscal restraint</td>
</tr>
<tr>
<td></td>
<td>Keynes liquidity preference theory</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neoclassical synthesis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Market discipline hypothesis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II ir → def</td>
<td>Post-Keynesian ‘horizontalists’</td>
<td>tight monetary policies contribute to deteriorate the fiscal situation</td>
<td>monetary authorities must assume responsibility for fiscal problems</td>
</tr>
<tr>
<td>III def → ir</td>
<td>Ricardian equivalence proposition</td>
<td>self-sustained cycle of deficit and interest rate increases</td>
<td>possible need for coordination of fiscal and monetary policies</td>
</tr>
<tr>
<td></td>
<td>Financial markets integration</td>
<td>it is indifferent to finance expenditures by taxes or debt</td>
<td>other variables are more relevant in the determination of interest rates and deficits</td>
</tr>
</tbody>
</table>

23 Also, the single-equation representations of the interest rate depend on the researcher's subjective pre-choice of the background theoretical model, which indicates the variables to be included in the estimation. Knot (1996) presents a more extensive enumeration of the merits and disadvantages of using these models in an investigation of the relation deficits/interest rates.
The procedure followed by most empirical studies, exclusively based on correlation analysis, is not able to distinguish between the three first hypothesis outlined above, always concluding in favour of I. Furthermore, if III is the true hypothesis, the studies which consider the deficit as an exogenous variable would be affected by problems of simultaneous-equation bias, producing biased and inconsistent estimates.

Additionally, it is also necessary to acknowledge the possibility of finding a negative relation between deficits and interest rates, leading to three additional hypothesis to consider. In fact, several studies report evidence of a negative relationship. Evans (1985) and Kolluri and Giannaros (1987), for example, claim that increasing deficits are associated with falling interest rates in the US. Evans (1987) and Boothe and Reid (1989) present evidence of a negative relation between debt and interest rates in, respectively, the US and Canada.

The criticisms on the standard econometric procedures led to the introduction of alternative approaches and econometric methods. The methodology adopted here, following the consistent approach of previous chapters, was initially suggested by Sims (1980) and it is based on the estimation of a vector autoregressive (VAR) model, where no *a-priori* judgement is made concerning the choice of the dependent variables or the dynamic structure of the model. Theory only influences the modelling strategy by suggesting which variables to include in the VAR.

One advantage of the VAR model, particularly useful for the objectives of this chapter, is the possibility of testing Granger non-causality. In view of the different empirical approaches available, and also considering the data restrictions, this seems the most adequate approach to study the relationship between the government’s total deficit or debt and the interest rate.

The evolution of the econometric procedures designed to test for non-causality were already extensively presented in chapter 3. The methodology used in each case depends on the order of integration of the variables involved and whether they are cointegrated or

---

24 Miller and Russek (1996) compare three alternative econometric methods which have evolved as a reaction against the methodology based on the Cowles Commission approach: 'Bayesian' (from the pioneer work of E. Learner), 'LSE methodology' (identified with D. Hendry) and 'VAR' (C. Sims). Applying these three methodologies to analyse the connection between deficits and interest rates, they conclude that the empirical results are sensitive to the econometric method preferred.
not. With nonstationary but cointegrated variables, the non-causality tests are performed in the context of the VECM. In all other cases, the tests are performed in the VAR, with the nonstationary variables pre-converted into a stationary process by the precise degree of differencing suggested by the ADF tests. When there is cointegration, estimation of an unrestricted VAR in first-differences is misspecified because it omits the equilibrium errors (see, for example, Campbell and Perron, 1991).

These more standard tests are complemented by the more recent LA-VAR methodology, which presents the main advantage of not requiring a priori knowledge of the order of integration/cointegration of the variables involved, which is particularly useful given the results of the unit root tests of Table 6.1. However, it is convenient to remember that the overall conclusions from the LA-VAR procedure and the standard tests may not coincide exactly in all occasions, which may be due to the lower power of the former type of non-causality tests.

6.5. The relation deficits - interest rates

Panel A of Table 6.3 presents the synthesised results of the tests on the causal relationship between the central government’s total surplus and the nominal interest rate. The table displays the order of integration of the variables involved in the model, the dummy variables necessary to induce well-behaved residuals, the results of the appropriate non-causality tests, and the lag-order of each model. When there is no unanimous result concerning the order of integration of a variable, both possibilities are considered, and non-coincident results are reported.

With the only exception of the United Kingdom, the tests indicate a strong and positive causal relationship running from the nominal interest rate to the total deficit of the central government. This suggests that, contrary to the ‘standard’ economic theory, it is usually the nominal interest rate that affects the government’s deficit and not the contrary. This relation reflects the importance of the interest expenditures on the government’s deficit, and depends on the amount of outstanding debt.
Table 6.3: Central government total surplus and the interest rates

<table>
<thead>
<tr>
<th>Order of integration:</th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ts ( i / r^{CPI} )</td>
<td>I(1) / I(0)</td>
<td>I(1) / I(0)</td>
<td>I(0) / I(1)</td>
<td>I(1) / I(0)</td>
<td>I(1) / I(1)</td>
<td>I(1) / I(0)</td>
</tr>
<tr>
<td>A) total surplus and nominal interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummies(^1) 1980,81 1980,81 1953,81 1981 1974</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-causality tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VECM(^4) i (\rightarrow) ts</td>
<td>30.8727**</td>
<td>n.c.</td>
<td>15.0226**</td>
<td>n.c.</td>
<td>n.c.</td>
<td></td>
</tr>
<tr>
<td>ts (\rightarrow) i</td>
<td>0.3913</td>
<td>23.8820**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>order 1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR(^4) i (\rightarrow) ts</td>
<td>6.8067**</td>
<td>6.9819**</td>
<td>6.5272**</td>
<td>0.0018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ts (\rightarrow) i</td>
<td>1.2195</td>
<td>0.1540</td>
<td>0.4619</td>
<td>0.7863</td>
<td></td>
<td></td>
</tr>
<tr>
<td>order 1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA-VAR i (\rightarrow) ts</td>
<td>7.4381**</td>
<td>5.4712**</td>
<td>13.6119**</td>
<td>6.6420**</td>
<td>0.1205</td>
<td></td>
</tr>
<tr>
<td>ts (\rightarrow) i</td>
<td>0.5485</td>
<td>1.4015</td>
<td>17.3561***</td>
<td>0.0803</td>
<td>0.2931</td>
<td></td>
</tr>
<tr>
<td>order 2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>B) total surplus and real interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummies(^1) 1952,58, 1953,58 1953 1973,1974 1975</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-causality tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VECM(^4) i (\rightarrow) ts</td>
<td>0.0023</td>
<td>0.3403</td>
<td>15.4740***</td>
<td>1.3696</td>
<td>0.9376</td>
<td>0.1585</td>
</tr>
<tr>
<td>ts (\rightarrow) i</td>
<td>0.1868</td>
<td>0.0005</td>
<td>1.5006</td>
<td>6.4674**</td>
<td>0.5708</td>
<td>0.4254</td>
</tr>
<tr>
<td>order 1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR(^4) i (\rightarrow) ts</td>
<td>0.6582</td>
<td>0.73109</td>
<td>8.5477**</td>
<td>4.6628</td>
<td>0.2913</td>
<td>0.01634</td>
</tr>
<tr>
<td>ts (\rightarrow) i</td>
<td>0.2167</td>
<td>1.0653</td>
<td>0.1938</td>
<td>9.9546**</td>
<td>6.6442</td>
<td>0.2800</td>
</tr>
<tr>
<td>order 2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

\(^1\) Intervention dummies are included in the model whenever necessary to whiten the residuals. \(^2\) Johansen cointegration test based on a model with a restricted intercept and no trend (n.c. indicates not cointegrated). \(^3\) First-differencing the I(1) variables. \(^4\) Based on Newey-West adjusted S.E.'s Bartlett weights, truncation lag=3. The hyphens (---) indicate that the test was not performed given the order of integration/cointegration of the variables. The asterisks (*), (**), and (***), indicate that the null hypothesis of non-causality was rejected at the, respectively, 10%, 5% and 1% levels of significance. In square brackets are the p-values of the statistics.

Illustrating this is the fact that a long-term equilibrium relationship was only found in Belgium and Italy. These are the only two countries where the value of the debt exceeds...
the present value of each respective GDP, leaving the deficit highly vulnerable to interest rate changes. In Belgium, for example, "a 1 percentage point drop in the average or effective interest rate paid on government debt translates into a decrease in the budget deficit of 1.3 per cent of GDP" (OECD Economic Surveys: Belgium, 1997, p. 43). In Italy, this same point drop in the interest rate, maintained over the 1995-2000 period would by itself "reduce the debt-GDP ratio directly by almost 5 points by the year 2000" (OECD Economic Surveys: Italy, 1997, p. 66).25

The effect of the deficits on the nominal interest rate appears only in the case of Italy, where a bidirectional and positive causal relationship was found between these two variables. It indicates the existence in this country of a self-fulfilling continuous cycle of nominal interest rate increases and budget deterioration.26 This is consistent with the clear indication of unsustainable fiscal policies in Italy. As for the other countries, these results suggest that in spite of evidence of unsustainable fiscal positions in most of them, this does not seem to create pressures on the interest rate.

Panel B of Table 6.3 examines this relationship but using the long-term ex post real interest rate. One first important observation is that the variables in the model usually present different orders of integration, restricting the choice of non-causality tests available. This can also be considered a prior, although not sufficient, indication of a possible lack of relationship between the variables.

In fact, the only two significant effects found, at the usual levels of significance, suggest a negative relation.27 In Germany the real rate has a negative effect on the level of the total deficit. The inverse effect, from the deficit to the real interest rate, appears in Italy (and the Netherlands, but it is a very weak effect) but also seems to be negative, as can be more informally confirmed in Figure A.10, for example.28

---

25 These OECD values are very close to the results obtained here. The long-term relation between the variables obtained in Table 6.3, quantified by the error correction term, was (standard errors in parenthesis) τr = -1.1481 i + 0.0394 for Belgium, and τr = -1.3213 i + 0.0538 for Italy. (0.1912) (0.0143) (0.1846) (0.0183) 26 The possibility of a vicious cycle in the Italian budget process was suggested for example by Fassio (1994: p. 5). 27 The lack of empirical evidence on the effect of the deficits on the real interest rate, contrary to what the 'standard' economic theory would suggest, has been documented in several empirical papers. A non-comprehensive list includes Boothe and Reid (1989) and Smithin (1994) for Canada, Makin (1983), Evans (1985), Kolluri and Giannaros (1987) and Findlay (1990) for the United States. 28 The correlation coefficient between the two Italian variables is -0.04.
The negative impact of the deficit on the real interest rate in Italy could be, for example, a consequence of a monetization policy. If the government reacts to a deterioration of the deficit by increasing the seigniorage revenues, the monetary easing will probably reduce the nominal interest rate and/or increase the inflation rate, and therefore the real interest rate tends to decrease. However, as seen in the previous chapter, although there is some weak evidence that inflation might have played a part in the government’s revenue policy, it was not possible to reject the hypothesis that deficits do not have a positive effect in seigniorage or inflation.

These results obtained with the real interest rate may also be due possibly to the non-verification of the Fisher effect. The post-war data may be dominated by the sharp negative relationship between nominal and real interest rates from about early 1970s. Higher inflation produced lower real rates in the 1970s, and then falling inflation caused higher real interest rates in late 1980s and 1990s.

As shown in the table, it was necessary in most tests to include intervention dummies in the model, in order to solve problems of especially non-normality of the residuals due to excess kurtosis, caused mainly by the high volatility of the variables in the beginning of the sample and the structural changes in the mid-1970s and beginning of the 1980s. Alternatively, and following the suggestion of Johansen and Juselius (1992), the present and one period lagged change in the price of oil ($\Delta p_{oi}$ and $\Delta p_{oi-1}$) were included in the model. This reduces considerably the problems of non-normality, but it does not eliminate them completely. It is able to capture the changes of the 1970s, for example, but it does not account for the structural change which took place in the beginning of the 1980s (presumably due to the introduction of restrictive monetary policies). To eliminate the problems of non-normality (and also some problems of serial correlation and heteroscedasticity) it is necessary to increase the number of lags beyond those suggested by the usual selection criteria, or to incorporate a dummy in the model. This is precisely what the recourse to this procedure intended to avoid. Nevertheless, the final results from this model including the changes in the price of oil do not contradict the conclusions previously obtained. The nominal interest rate affects the government’s total deficit in all countries (even in the UK now) while the opposite effect only occurs in Italy.
Finally, in the sub-sample [1972-96], the tests with the nominal interest rate corroborate all the conclusions achieved with the full sample: except for the Netherlands and the United Kingdom, in all other countries the nominal interest rate has a positive effect on the central government's total deficit while the opposite effect does not seem to be significant. When the real interest rate is used, the only significant results concern Germany (possibly due to the importance of the effects of the reunification on this sub-sample), where the test suggests a bidirectional, negative, relation between the deficit and the real interest rate. For all other countries, the deficit does not seem to have any effect on the real interest rate and vice-versa.

6.6. An analysis of sovereign risk

While the previous section analysed the effect on interest rates of the government's deficit, a flow variable, this section studies the influence on a stock fiscal variable, the accumulated government debt. Although, theoretically, debt is simply an accumulation of successive past deficits, each variable allows the investigation of different relations with the interest rates. The analysis of the relationship between debt and interest rates raises for example questions of sovereign risk. Do the markets react to different fiscal positions by including a risk premium on the interest rate demanded on government bonds? The relationship was again tested separately for the nominal and the real interest rates.

The tests in panel A of Table 6.4 found a generalised negative relationship between the level of debt and the nominal interest rate. This result, although somewhat unexpected in theoretical grounds, confirms the overall impression given more directly by a visual inspection of Figure A.10, for example.
Table 6.4: Central government debt and the interest rates

<table>
<thead>
<tr>
<th>Order of integration:</th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d ) / ( r^{\text{CPI}} )</td>
<td>I(1) / I(0)</td>
<td>I(1) / I(0)</td>
<td>I(0) / I(1)</td>
<td>I(1) / I(0)</td>
<td>I(1) / I(1)</td>
<td>I(1) / I(0)</td>
</tr>
<tr>
<td>A) Debt and nominal interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummies ( t )</td>
<td>1993</td>
<td>1979</td>
<td>1953</td>
<td>1981</td>
<td>1954</td>
<td>1964,74</td>
</tr>
<tr>
<td>Non-causality tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VECM ( t )</td>
<td>i ( \rightarrow ) d</td>
<td>41.8242 ( ** )</td>
<td>21.5264 ( ** )</td>
<td>---</td>
<td>n.c.</td>
<td>18.3975 ( ** )</td>
</tr>
<tr>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .045 ]</td>
<td>[ .000 ]</td>
<td>[ .007 ]</td>
<td>[ .007 ]</td>
</tr>
<tr>
<td>d ( \rightarrow ) i</td>
<td>24.0897 ( *** )</td>
<td>19.2451 ( *** )</td>
<td>13.1159 ( *** )</td>
<td>9.9286 ( *** )</td>
<td>[ .004 ]</td>
<td>[ .007 ]</td>
</tr>
<tr>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
</tr>
<tr>
<td>order</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>VAR ( t )</td>
<td>i ( \rightarrow ) d</td>
<td>---</td>
<td>---</td>
<td>8.1236 ( ** )</td>
<td>7.8581 ( ** )</td>
<td>---</td>
</tr>
<tr>
<td>[ .044 ]</td>
<td>[ .020 ]</td>
<td>[ .000 ]</td>
<td>[ .034 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
</tr>
<tr>
<td>d ( \rightarrow ) i</td>
<td>0.5017</td>
<td>6.7906 ( ** )</td>
<td>[ .919 ]</td>
<td>[ .034 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
</tr>
<tr>
<td>order</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>LA-VAR ( t )</td>
<td>i ( \rightarrow ) d</td>
<td>20.4282 ( ** )</td>
<td>8.4480 ( ** )</td>
<td>2.5391</td>
<td>2.5131</td>
<td>11.5867 ( ** )</td>
</tr>
<tr>
<td>[ .000 ]</td>
<td>[ .038 ]</td>
<td>[ .028 ]</td>
<td>[ .028 ]</td>
<td>[ .009 ]</td>
<td>[ .009 ]</td>
<td>[ .009 ]</td>
</tr>
<tr>
<td>d ( \rightarrow ) i</td>
<td>6.5702 ( * )</td>
<td>5.7501</td>
<td>.095687</td>
<td>4.7323</td>
<td>8.7289 ( ** )</td>
<td>6.9246 ( ** )</td>
</tr>
<tr>
<td>[ .087 ]</td>
<td>[ .124 ]</td>
<td>[ .953 ]</td>
<td>[ .094 ]</td>
<td>[ .033 ]</td>
<td>[ .031 ]</td>
<td>[ .031 ]</td>
</tr>
<tr>
<td>order</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>B) Debt and real interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummies ( t )</td>
<td>1952,74</td>
<td>1953,58,79</td>
<td>1953</td>
<td></td>
<td></td>
<td>1964,75</td>
</tr>
<tr>
<td>Non-causality tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VECM ( t )</td>
<td>( r^{\text{CPI}} ) ( \rightarrow ) d</td>
<td>---</td>
<td>---</td>
<td>12.2267 ( ** )</td>
<td>n.c.</td>
<td>6.9140 ( * )</td>
</tr>
<tr>
<td>[ .002 ]</td>
<td>[ .075 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
<td>[ .047 ]</td>
<td>[ .947 ]</td>
<td>[ .947 ]</td>
</tr>
<tr>
<td>d ( \rightarrow ) ( r^{\text{CPI}} )</td>
<td>23.9493 ( ** )</td>
<td>8.9217 ( ** )</td>
<td>[ .000 ]</td>
<td>[ .030 ]</td>
<td>[ .000 ]</td>
<td>[ .000 ]</td>
</tr>
<tr>
<td>order</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>VAR ( t )</td>
<td>( r^{\text{CPI}} ) ( \rightarrow ) d</td>
<td>0.1591</td>
<td>0.4909</td>
<td>---</td>
<td>1.4715</td>
<td>---</td>
</tr>
<tr>
<td>[ .690 ]</td>
<td>[ .484 ]</td>
<td>[ .479 ]</td>
<td>[ .947 ]</td>
<td>[ .947 ]</td>
<td>[ .947 ]</td>
<td>[ .947 ]</td>
</tr>
<tr>
<td>d ( \rightarrow ) ( r^{\text{CPI}} )</td>
<td>4.0486 ( * )</td>
<td>2.0383</td>
<td>1.9828</td>
<td>7.6198 ( ** )</td>
<td>[ .044 ]</td>
<td>[ .022 ]</td>
</tr>
<tr>
<td>[ .013 ]</td>
<td>[ .371 ]</td>
<td>[ .002 ]</td>
<td>[ .002 ]</td>
<td>[ .002 ]</td>
<td>[ .002 ]</td>
<td>[ .002 ]</td>
</tr>
<tr>
<td>order</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>LA-VAR ( t )</td>
<td>( r^{\text{CPI}} ) ( \rightarrow ) d</td>
<td>4.9716 ( * )</td>
<td>.75363</td>
<td>4.0592</td>
<td>1.8090</td>
<td>0.8864</td>
</tr>
<tr>
<td>[ .083 ]</td>
<td>[ .385 ]</td>
<td>[ .131 ]</td>
<td>[ .179 ]</td>
<td>[ .829 ]</td>
<td>[ .879 ]</td>
<td>[ .879 ]</td>
</tr>
<tr>
<td>d ( \rightarrow ) ( r^{\text{CPI}} )</td>
<td>7.9845 ( ** )</td>
<td>0.7186</td>
<td>3.2332</td>
<td>0.0205</td>
<td>16.4037 ( ** )</td>
<td>6.7275</td>
</tr>
<tr>
<td>[ .018 ]</td>
<td>[ .397 ]</td>
<td>[ .199 ]</td>
<td>[ .886 ]</td>
<td>[ .001 ]</td>
<td>[ .009 ]</td>
<td>[ .009 ]</td>
</tr>
<tr>
<td>order</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

\(^1\) Intervention dummies are included in the model whenever necessary to whiten the residuals. \(^2\) Johansen cointegration test based on a model with an unrestricted intercept and a restricted trend (restricted intercept and no trend for the UK in A and for N in B), (n.c. indicates not cointegrated). \(^3\) First-differencing the I(1) variables. \(^4\) Based on Newey-West adjusted S.E.’s Bartlett weights, truncation lag=3. The hyphens (---) indicate that the test was not performed given the order of integration/cointegration of the variables. The asterisks (*), (**), and (***), indicate that the null hypothesis of non-causality was rejected at the, respectively, 10%, 5% and 1% levels of significance. In square brackets are the p-values of the statistics.
One possible explanation for this behaviour concerns the role of uncertainty about future taxation, and was pointed out in 1974 by Barro (see also Seater, 1993). For example, if the prudent individuals perceive the current deficit and the resulting accumulated debt as a future liability, as assumed by the RET, but there is uncertainty about the future tax burden they will support, the increase in precautionary private savings may more than compensate the reduction in public savings, with a negative overall effect in the interest rate. Becker (1997), for example, found empirical support for this hypothesis. Using US data for the period 1960-1993, he shows that private individuals discount more than 100% of the future tax requirements. This finding contradicts the standard macroeconomic theory, which anticipates zero or at least less than 100% discounting (see, for example, Patinkin, 1965: p. 289), but it is also contrary to the RET, which anticipates exactly 100% tax discounting. 29

An alternative explanation, still within the RET theoretical framework, was advanced by Boskin (1978) and Evans (1985). They point out that a deficit caused by a tax reduction implies a higher after-tax real rate of return on savings. The additional growth in private savings, together with the increase necessary to account for the postponed taxes, will more than compensate for the government’s dissaving. The final result suggests a possible negative relation between deficits/debt accumulation and the interest rate.

Panel B in Table 6.4 examines whether the same conclusions can be extended to the relation between debt and the real interest rate, and rejects this hypothesis. The results of the non-causality tests reveal that, in general, the level of the government debt has a strong positive influence on the real interest rate. This effect was not found only in the cases of France and Italy. There is also some evidence of a bidirectional relationship in Belgium, Germany and the Netherlands, (although only in Germany it is significant at the 5% level of significance). 30

29 Using a reduced-form equation derived from an IS/LM model, Knot (1996) has also found evidence of a negative relationship between the nominal interest rate and public debt (in, among others, France, Italy, Netherlands and UK, but not in Germany). He explains this negative relation as showing that the private individuals have discounted at least part of their future tax liabilities (to pay for the debt) in forming their consumption plans.

30 The results in panel B are particularly sensitive to the deterministic elements included in the VAR, the model chosen for the Johansen cointegration test, the dummies included and the number of lags and, therefore, they should be interpreted with extra caution. They are also sensitive to the time period considered. For the sub-sample [1972-96], for example, the tests do not present evidence of a positive effect of the debt on the real interest rate.
The absence of statistically significant results for France and especially for Italy, one of the EU's largest debtors, are somewhat unexpected. One possible explanation is the fact that these two countries have maintained capital controls for a longer period than any other country in the sample (see, for example, Holmes and Luintel, 1999). This protection of the internal financial markets, notably in the case of Italy, may have allowed the government to influence the rate at which it obtains financing from the market.31

6.7. General government fiscal variables and the interest rate

This section will examine whether the results obtained above are affected by using a different, broader level of the government accounts. The general government accounts include the central, local and regional governments, and the social security funds. The exact results of the unit root tests on the fiscal variables can be found in Table 4.11, and will not be reproduced here. For Belgium and Italy, for whom there is now a very slightly shorter sample, the unit root tests (not shown) were repeated for the interest rate variables.32 Table 6.5 presents the results of the non-causality tests.

With the exception of Italy and the Netherlands, the tests of non-causality suggest a strong and positive causal effect running from the nominal interest rate to the general government total deficit.33 The inverse effect also appears, but only in the cases of Belgium and Italy, the two most indebted countries in the group.34

The sporadic differences in results when using central or general government data highlight particularly the importance of the social security funds. In Belgium, for example, there is now a causal effect from the deficits to the nominal interest rate which

31 Some authors (e.g. Pivetti, 1998) even propose the maintenance of some capital controls as a necessary measure to avoid these fiscal effects on the interest rate.
32 Recall that for Belgium the sample period is 1953-1996 while for Italy is 1951-1996.
33 In Italy, "the central government is the chief beneficiary of lower interest rates" (OECD Economic Surveys: Italy, 1996, p. 41).
34 The only significant results (not shown) of the non-causality tests between the general government's total deficit and the real interest rate suggest that, in Germany and Italy, the latter has a negative influence on the former. No other statistically significant effect was found in any country, confirming the lack of results previously obtained with the central government's data.
did not appear with the central government’s data. In fact, Belgium is the only country where on average for the entire period considered here, the deficits of the general government have been higher than those of the central government.

Table 6.5: General government total surplus and nominal interest rate

<table>
<thead>
<tr>
<th>Order of integration:</th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ts(^G)</td>
<td>l(1)</td>
<td>l(1)</td>
<td>l(1)</td>
<td>l(1)</td>
<td>l(0)</td>
<td>l(1)</td>
</tr>
<tr>
<td>i</td>
<td>l(1)</td>
<td>l(1)</td>
<td>l(0)</td>
<td>l(1)</td>
<td>l(1)</td>
<td>l(1)</td>
</tr>
<tr>
<td>Dummies(^f)</td>
<td>1980,81</td>
<td>1953</td>
<td>1975,81</td>
<td>1974</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Non-causality tests:

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>VECM(^i) ts(^G) (\rightarrow) i</td>
<td>28.1646***</td>
<td>2.4395</td>
<td>10.6584***</td>
<td>22.7983***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[.000]</td>
<td>[.295]</td>
<td>[.005]</td>
<td>[.000]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ts(^G) (\rightarrow) i</td>
<td>6.1966**</td>
<td>8.8960**</td>
<td>0.1420</td>
<td>7.7556**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR(^i) ts(^G) (\rightarrow) i</td>
<td>2.5721</td>
<td>5.2889*</td>
<td>0.3350</td>
<td>1.7363</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[.004]</td>
<td>[.964]</td>
<td>[.831]</td>
<td>[.852]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This may justify the evidence of a self-sustained cycle of deficits and interest rate increases in Belgium, which previously have only appeared in Italy, with central government data. Using precisely the general government data, the *OECD Economic Surveys: Belgium* (1992: p. 29) warned that in Belgium “the snowball effect - *i.e.,* the self-sustaining increase in the debt/GNP ratio as a result of interest payments - which had been arrested in 1989, resumed”.

---

\(^i\)Intervention dummies are included in the model whenever necessary to whiten the residuals. \(^j\)Johansen cointegration test based on a model with a restricted intercept and no trend (n.c. indicates not cointegrated).

\(^f\)First-differencing the l(1) variables. \(^g\)Based on Newey-West adjusted S.E.'s Bartlett weights, truncation lag=3. The hyphens (---) indicate that the test was not performed given the order of integration/cointegration of the variables. The asterisks (*), (**), and (***) indicate that the null hypothesis of non-causality was rejected at the, respectively, 10%, 5% and 1% levels of significance. In square brackets are the p-values of the statistics.
All the tests performed so far do not reveal whether the relations found between the fiscal variables and the interest rates are the result of domestic fundamentals in each country, or if there are common external influences concealing or emphasising the influence of these fundamentals. The next two sections will try to distinguish these effects by analysing the interest rate differentials, which remove the external influences, and the aggregate variables, which focus exclusively on those external determinants.

6.8. The interest rate differentials: fiscal discipline through the financial markets

By examining the interest rate differential to some average or third country interest rate, and comparing the results with those obtained with the domestic interest rate, it is possible to observe the extent to which external factors constrain the effects of internal influences on the domestic rate. Those external influences are typically proxied for the EU countries by German (Craig, 1991 and 1994, and Knot, 1998) or US interest rates (Ibrahim and Kumah, 1996, for example).

One interesting possibility, analysed by several authors, is the hypothesis that financial markets may discipline national government's fiscal policy. Interest rates on public bonds of less disciplined governments would increase relatively to the others. This situation can be more easily depicted by dividing the causes of an interest rate differential, with perfect capital mobility, into

\[ i - i^* = \Delta r^* + fcrp + drp , \]  

where the left-hand side is the difference between the domestic and the foreign long-term interest rates on government bonds with identical maturity, liquidity and tax treatment.

---

35 Besides this effect on interest rate differentials, financial markets may also discipline governments through the 'depreciation threat' (see for example Heinemann, 1998).

36 Alternatively, a fourth term could have been added to the right-hand side of equation (6.4), representing the country-premium associated with the barriers to capital mobility. These barriers have been completely removed in different dates by the countries included in the sample, starting with the UK in 1979 and ending with France and Italy only in 1990.
\( \Delta er^p \) is the exchange rate depreciation expectation, \( ferp \) represents the foreign exchange risk premium, and \( drp \) is the default risk premium.

Equation (6.4) allows a better understanding of the way interest differentials react to the fiscal position of a government under different exchange rate systems and how things will change in the particular case of the European Union, with EMU. The exchange rate related determinants of the interest rate, the first two elements on the right-hand side, depend on the exchange rate system adopted and will disappear within a monetary union. The default risk premium depends basically on the sustainability of government policies and will remain, with perfect capital mobility, as the only determinant of the interest rate differentials. Some authors even claim that the disappearance of the other two factors may not lead necessarily to a narrowing of the interest rate differentials because the default risk premium may increase in a monetary union. This premium will be larger the more credible is the no-bailout rule of the Maastricht Treaty, prohibiting the European Central Bank and the national central banks from buying public debt directly from the issuer.\(^{37}\)

The interest rate used in this sub-section is the difference between the domestic rate and the US (until 1971) or the German (after 1972) rates. The choice of these two reference countries in these two particular time periods is motivated by their 'anchor' functions in, respectively, the Bretton Woods fixed exchange rate system and the European Monetary System. For Germany, the differential is with the US rate for the whole period.

With the exception of the Netherlands, all nominal interest rate differentials were found to be nonstationary (Table 6.1). The empirical literature is not unanimous on this matter. Using smaller samples for roughly the same group of countries, Knot (1998) for example claims they are stationary (sample 1981-92) while Katsimbris and Miller (1993) report the contrary (sample 1979-88).

The only significant result of the tests suggests a bidirectional positive relationship between the Italian total deficit and its nominal interest rate deviation from the US and German rates, confirming the previous results with the interest rate in levels. It suggests that an increase in the Italian's government deficit will cause its nominal interest rate to

---

\(^{37}\) Art. 21 of the Protocol on the European System of Central Banks.
increase relatively to the foreign rate which, in turn, will further deteriorate the fiscal balance. Italy seems to be the only country where domestic fiscal determinants play an important role in defining the nominal interest rate.  

Table 6.6: Central government total surplus and nominal interest rate differential

<table>
<thead>
<tr>
<th>Order of integration:</th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ts</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(0)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
</tr>
<tr>
<td>id</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(0)</td>
<td>I(1)</td>
</tr>
<tr>
<td>dummies</td>
<td>1981</td>
<td>1953.75</td>
<td>1972</td>
<td>1974</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Non-causality tests:

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>N</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>VECM*</td>
<td>n.c.</td>
<td>n.c.</td>
<td>---</td>
<td>n.c.</td>
<td>---</td>
<td>n.c.</td>
</tr>
<tr>
<td>VAR*</td>
<td>i^d \rightarrow ts</td>
<td>2.1525</td>
<td>2.6998*</td>
<td>0.0727</td>
<td>6.6669***</td>
<td>2.6167</td>
</tr>
<tr>
<td></td>
<td>ts \rightarrow i^d</td>
<td>0.3503</td>
<td>0.6595</td>
<td>0.3929</td>
<td>4.5664**</td>
<td>5.6770*</td>
</tr>
<tr>
<td>order</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>LA-VAR*</td>
<td>i^d \rightarrow ts</td>
<td>3.0088*</td>
<td>2.0017</td>
<td>1.3893</td>
<td>3.7642</td>
<td>2.3085</td>
</tr>
<tr>
<td></td>
<td>ts \rightarrow i^d</td>
<td>0.8555</td>
<td>0.9089</td>
<td>1.9419</td>
<td>4.1419</td>
<td>0.6052</td>
</tr>
<tr>
<td>order</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

*Intervention dummies are included in the model whenever necessary to whiten the residuals. *Johansen cointegration test based on a model with a restricted intercept and no trend (n.c. indicates not cointegrated). *First-differencing the I(1) variables. *Based on Newey-West adjusted S.E.'s Bartlett weights, truncation lag=3. The hyphens (---) indicate that the test was not performed given the order of integration/cointegration of the variables. The asterisks (*), (**), and (***) indicate that the null hypothesis of non-causality was rejected at the, respectively, 10%, 5% and 1% levels of significance. In square brackets are the p-values of the statistics.

One other interesting result from these tests is that while using the domestic individual interest rates it was found that they have a strong positive impact on the deficit, with interest rate differentials that causal effect is only statistically significant (at the usual levels) in the case of Italy. This suggests that for all other countries this effect is

38 In fact, Italy is the only country where the interest rate differentials with Germany have remained substantial in the last decades. Craig (1994) have also found this effect for Italy.

39 The effect is also noticeable in Belgium and France, but the hypothesis of non-causality can only be rejected at the 10% level of significance. Given the similarities between the Italian and Belgian fiscal situation, it would be reasonable to expect a similar result. Both countries share a history of high deficits and debt since the early 1970s, well above any of the other countries in the sample. However, their interest rate differentials with Germany are significantly apart. Artus (1998), for example, explains this difference with the existence of multiple equilibria. Although both countries present a poor fiscal record, Belgium has
primarily driven either by external factors or by similar movements in the domestic interest rates. The next section examines this hypothesis more directly.

Using a smaller sample and monthly data (1987.01-1993.12) for a similar group of countries, Craig (1994) found that only in Belgium did the deficit affect the interest rate differential with Germany. He uses general government data from the OECD, which confirms the importance of the social security funds and the regional authorities in this country, as previously found in section 6.7. Ibraim and Kumah (1996), using the short-term interest rate differential with the US rate, also conclude that variations in the differential in Germany and the United Kingdom, among other countries, are explained more by monetary than by fiscal innovations. Knot (1998) reports mixed conclusions, with the long-run differentials being usually affected by the deficits and short-term interest differentials not affected.

When considering exclusively the second half of the sample (1972-96), the results (not shown) continue to indicate a positive causal effect running from the interest rate differential to the total deficit in Belgium and Italy, the two most indebted members of the EU.\(^40\) The results of the same test but with the real interest rate differentials do not reveal any statistically significant positive relationship, and therefore will also not be presented here.

Overall, the evidence presented in this section does not generally support the market discipline hypothesis. With the exception of Italy, and possibly Belgium when using general governments' data, there is no evidence suggesting that investors differentiate between different fiscal positions by including different risk premia on the interest rates of bonds placed by the governments.

---

\(^{40}\) When considering this restricted sample, the interest rate differential is measured exclusively with reference to the German rates. For Germany the reference is the US rate.
6.9. European-wide effects

As noticed from the observation of Figure A.10, for example, the interest rates in the European Union have followed a similar general pattern across countries, both in nominal and real terms. The nominal interest rate exhibits a double peak, with significant increases in the aftermath of both oil crises, and a third less pronounced peak in the beginning of the nineties, probably due to the impact of the German unification process. The real interest rate has fallen suddenly and almost simultaneously in all EU countries in the mid-1970s, reaching in general negative values, recovering in the early 1980s to stabilise in the last decade at levels above those prior to the 1970s.

Christiansen and Pigott (1997) found evidence of an increase in the covariation of long-term interest rates among some EU countries in the 1990s, implying that interest rates became possibly less subject to the influence of domestic determinants. Using panel data techniques, Gagnon and Unferth (1995) found a high correlation of *ex post* real interest rates across OECD countries for the period 1977-93.

These common trends in the evolution of the interest rates and also, to a certain extent, of the fiscal variables (see Holmes and Luintel, 1999), suggest the importance of analysing the common factors in the relationship between these variables. Several authors have suggested that, in a market with high capital mobility, it is the aggregate fiscal variables which influence the interest rates. The growing integration of world financial markets allows governments to finance their deficits through external borrowing, possibly spreading the effects of deficits and reducing its domestic consequences. However, collective deficits and debt may still have an impact on international interest rates.

While in the previous section the intention was to examine exclusively the influence of internal factors acting through the risk premia, removing all the external influences, this section is concerned only with those common elements in the determination of the

---

41 The historically high levels in the early eighties were mainly due to the disinflationary policies pursued in most countries.
42 The negative values of the mid seventies may be explained by the strict financial regulations which prevented the nominal interest rate to fully reflect the high inflation rates prevailing.
interest rates. Disregarding the different fiscal evolution of the individual countries, the tests concentrate on the explanation of its average development.

One other potential advantage of studying the aggregate relationship is the stability of the model. In the literature on aggregate money demand functions in the European Union, it has been argued that an aggregate money demand function is more stable than its individual country counterparts (see, for example, Knot, 1996, Cassard et al., 1997, and Angeloni et al., 1999).

This section will examine the relationship between the aggregate total deficit/debt of the six countries under analysis, in relation to GDP, and the average interest rate (Figure A.11 in the appendix). The fiscal variables of each country were converted into dollars using the average market exchange rate, aggregated, and then divided by the aggregate GDP (computed in the same way). The average interest rate was weighted by the proportion of each country's GDP on the total of the six countries for each year. 44

The results of the unit root tests on the series of the aggregated variables are shown on the first columns of Table 6.7. With the exception of the real interest rate, all other aggregated series seem to possess a unit root.

The tests displayed in panel A indicate that the aggregate deficit of the six European countries is affected by both the nominal and the real interest rates, while the opposite relation, suggested in the theoretical literature, was not found. This confirms the general results of the individual tests with the nominal rate. It also extends the same conclusions for the case of the real interest rate, for which the individual tests were not powerful enough to indicate any statistically significant result.

Using a reduced-form equation derived from an IS/LM model, Knot (1996) also found evidence of a positive relationship between the aggregate deficit and the nominal interest rate, in a group of five EU countries. However, he interprets it as evidence that the deficits affect the nominal interest rate, not the opposite. In contrast, Evans (1987) did not find any relationship between the aggregate deficit and the real interest rate of six OECD countries.

44 All the tests were also performed with simple averages of the ratio of the total surplus to GDP and the nominal and real interest rate, without affecting any conclusion. Nunes and Stemitsiotis (1995), for
### Table 6.7: Central government aggregated fiscal variables and interest rates

<table>
<thead>
<tr>
<th></th>
<th>ADF$_{SBC}$</th>
<th>ADF$_{ST}$</th>
<th>PV</th>
<th>dummies$^\dagger$</th>
<th>VECM$^\dagger$</th>
<th>VAR$^\dagger$</th>
<th>LA-VAR $^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A) Total surplus and interest rates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$i^t$</td>
<td>-2.0041</td>
<td>-1.5861</td>
<td>-3.0515</td>
<td>1975 $i^t \rightarrow ts^t$</td>
<td>20.2340'***</td>
<td>15.6045'***</td>
<td>29.3904'***</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td></td>
<td></td>
<td>[.000]</td>
<td>[.000]</td>
<td>[.000]</td>
</tr>
<tr>
<td>[1969]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ts^t$</td>
<td>-2.0250</td>
<td>-2.0250</td>
<td>-3.6225</td>
<td>1975 $ts^t \rightarrow i^t$</td>
<td>1.4912</td>
<td>0.0801</td>
<td>2.6586</td>
</tr>
<tr>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td></td>
<td></td>
<td>[.684]</td>
<td>[.961]</td>
<td>[.447]</td>
</tr>
<tr>
<td>[1974]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^t$</td>
<td>-3.6698'**</td>
<td>-3.6698'**</td>
<td>-5.2380'**</td>
<td>1975 $ts^t \rightarrow r^t$</td>
<td>0.2936</td>
<td>2.1197</td>
<td></td>
</tr>
<tr>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td></td>
<td></td>
<td>[.588]</td>
<td>[.145]</td>
<td></td>
</tr>
<tr>
<td>[1981]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B) Debt and interest rates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$i^t$</td>
<td>-2.0041</td>
<td>-1.5861</td>
<td>-3.0515</td>
<td>1981 $i^t \rightarrow d^t$</td>
<td>n.c.</td>
<td>1.6917</td>
<td>8.1695'**</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td></td>
<td></td>
<td>[.193]</td>
<td>[.017]</td>
<td></td>
</tr>
<tr>
<td>[1969]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d^t$</td>
<td>-1.2078</td>
<td>-1.8419</td>
<td>2.7418</td>
<td>1972 $d^t \rightarrow i^t$</td>
<td>0.4789</td>
<td>0.3567</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td></td>
<td></td>
<td>[.489]</td>
<td>[.837]</td>
<td></td>
</tr>
<tr>
<td>[1972]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^t$</td>
<td>-3.6698'**</td>
<td>-3.6698'**</td>
<td>-5.2380'**</td>
<td>1981 $d^t \rightarrow r^t$</td>
<td>0.0583</td>
<td>0.8850</td>
<td></td>
</tr>
<tr>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td></td>
<td></td>
<td>[.809]</td>
<td>[.347]</td>
<td></td>
</tr>
<tr>
<td>[1981]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^\dagger$See notes in Table 6.1. $^\dagger$Intervention dummies are included in the model whenever necessary to whiten the residuals. $^\dagger$Johansen cointegration test based on a model with a restricted intercept and no trend (unrestricted intercept and restricted trend in panel B), (n.c. indicates not cointegrated). $^\dagger$First-differencing the I(1) variables. $^\dagger$Based on Newey-West adjusted S.E.'s Bartlett weights, truncation lag= 3. The asterisks (*), (***) and (****) indicate that the null hypothesis of non-stationarity/non-causality was rejected at the, respectively, 10%, 5% and 1% level of significance. In square brackets are the $p$-values of the statistics.

Pursuing the same methodology, panel B of Table 6.7 presents the results of the tests with the aggregate stock of debt being used in place of the total deficit. The results show that, in aggregate terms, the nominal interest rate has a negative effect on the stock of public debt, while the inverse effect does not seem to exist. The tests also did not find a causal relationship, in any direction, between the aggregate debt and the aggregate real interest rate. Barro and Sala-i-Martin (1990) estimated a reduced-form equation of aggregated variables to explain the 'world' real interest rate and found also that the fiscal example, used as weights of the aggregated interest rates the share of the country in SDR composition, while the fiscal variables were weighted by the country's share in total GDP.
variables do not affect the interest rate. On the contrary, Ford and Laxton (1995), for example, found evidence that real interest rates are influenced by world debt levels.

6.10. Summary of the main results

This chapter has examined the long-run relationship between the fiscal variables and the interest rate. Several different, but complementary, samples and econometric techniques have been employed, with the objective of increasing the robustness of the results and explore this issue from distinct perspectives. Overall, the results have been reasonably consistent, and are summarily displayed in Table 6.8 and Table 6.9.

Table 6.8: The relation deficits-interest rates, summary results

<table>
<thead>
<tr>
<th>Central Government</th>
<th>Aggregate series</th>
<th>General Government</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual series</td>
<td>Differentials</td>
<td></td>
</tr>
<tr>
<td>def - i</td>
<td>$i \rightarrow \uparrow \text{def: B, F, G, I, N}$</td>
<td>$i^A \rightarrow \uparrow \text{def}^A$</td>
</tr>
<tr>
<td>def $\rightarrow \uparrow \text{i: I}$</td>
<td>$\text{def}^0 \rightarrow \uparrow \text{def: I}$</td>
<td>$\text{def} \rightarrow \uparrow i^0: I$</td>
</tr>
<tr>
<td>$\text{def} - r^{\text{CPI}}$</td>
<td>$r^{\text{CPI}} \rightarrow \downarrow \text{def: G}$</td>
<td>$r^A \rightarrow \uparrow \text{def}^A$</td>
</tr>
<tr>
<td>$\text{def} \rightarrow \downarrow r^{\text{CPI}}: I$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The countries are identified by the respective initials. The $\rightarrow$ indicates the direction of causality and $\uparrow$ a positive effect.

The first prominent result is the lack of empirical evidence showing a generalised positive impact of the deficits on the interest rate, contrary to what the conventional economic theory suggests. This result is largely supported in the empirical literature, and is usually justified with the RET. However, other explanations have more recently been advanced in the literature, reviewed in section 6.2.2, such as the integration of the financial markets, or the proposition that the interest rate is fundamentally a monetary and not a fiscal phenomenon.

On the contrary, the empirical evidence strongly suggests that it is generally the level of the interest rate which positively affects the level of the government’s deficit. There is
evidence of this effect in all the countries considered, as well as with the pooled variables. The interest rate affects the government's budget constraint through different channels, which may be divided according to the various components of the total deficit, 

\[ T_{\text{def}_i} = (G_i - T_i) + i_i D_{c,t} + \Delta M_i^N / P_t. \]

- primary deficit: a restrictive monetary policy, by increasing the interest rate, may reduce the levels of private investment, employment and output, with adverse effects on both government expenditures and revenues. On the one hand, the level of expenditures rises, impelled either by the automatic stabilisers or by discretionary policies intended to reduce the effects of the business cycle. On the other hand, the amount of taxes is reduced with the decline in economic activity. These effects are mainly due to the relative inertia of the primary surplus, which is very difficult to adjust with sufficient haste and flexibility, for economic, political and even institutional reasons.

- interest payments: tight monetary policies affect the costs of servicing the outstanding stock of debt in two ways. Directly, because of the higher interest rates. Indirectly, because if this policy is successful in reducing the inflation rate, it will curtail the erosion of the real value of debt.

- seigniorage: by reducing monetary expansion, a restrictive monetary policy diminishes this source of government revenues, which will have to be financed through debt.

When considering the real interest rate, the tests continue to present no evidence of an effect from deficits in any country. At the aggregate level there is evidence of the opposite effect, as happened with the nominal rate, although the tests in the individual countries could only find a negative relationship in two cases.

\[ \text{228} \]

\[ 45 \text{ Ultimately, the expectation of considerable deficit increases, caused by higher debt servicing costs, may push away the investors and jeopardise the financing perspectives of the government. See for example Alesina, Prati and Tabellini (1990) for a model of debt runs.} \]

\[ 46 \text{ Besides, if the high interest rates dampen economic activity, the lower rate of growth implies an increase in the debt-GDP ratio and a decrease in the change of base money. Dahan (1998) refers two additional effects which may arise in an open economy with perfect mobility of capital: the 'sterilisation effect', which contributes to a higher deficit due to the temporary gap between the domestic and foreign interest rate during the operation of sterilisation. With a restrictive monetary policy, the central bank sells domestic bonds at a lower price (because of the increase in the domestic interest rate) and buys foreign bonds at a higher price (because meanwhile the foreign interest rate is lower than the domestic rate); the 'swapping effect', which may offset the previous one, occurs if the government reduces its most expensive domestic debt and increases the less expensive foreign debt.} \]
Italy, with central, and Belgium, with general government data, are the only countries where a bidirectional relationship was found between the deficit and the nominal interest rate, suggesting the existence of a self-sustained cycle between these variables. Both countries share a history of very high deficits and debt since the early 1970s, well above any of the other countries in the sample. The fact that this bidirectional effect in Belgium appears only with general government data is consistent with the observation that, over the whole period, this is the only country where the average deficit of the general government is above the average deficit of the central government, emphasising the importance of the local authorities and, especially, the social security funds.

Table 6.9: The relation debt-interest rates, summary results

<table>
<thead>
<tr>
<th>Central Government</th>
<th>individual series</th>
<th>aggregate series</th>
</tr>
</thead>
<tbody>
<tr>
<td>d - i</td>
<td>i → ↓ d: B, F, G, I, N, UK</td>
<td>i → ↓ d^#</td>
</tr>
<tr>
<td>d → ↓ i: B, F, I, N, UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d - r^{CPI}</td>
<td>r^{CPI} → ↑ d: G</td>
<td>...</td>
</tr>
<tr>
<td>d → ↑ r^{CPI}: B, G, N, UK</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The countries are identified by the respective initials. The → indicates the direction of causality and ↓ a negative effect.

Table 6.9 summarises the results obtained when the stock of debt replaces the total deficit. Two distinct outcomes arise when using the nominal or the real interest rates. On the one hand, the evidence points overwhelmingly to a negative relationship between the level of debt and the nominal interest rate in all countries. This can also be inferred, more informally, from a visual inspection of Figure A.10.

On the other hand, the tests have detected a positive effect of the stock of debt on the real interest rate, in four of the six countries analysed. The exceptions, France and Italy, may be perhaps explained by the policy of capital controls maintained by these countries until the early 1990s, many years after all the others. This policy, designed to support the respective exchange rates, allowed these countries to artificially maintain lower real interest rates on the public bonds.

47 Over the whole sample, the average central government's deficit-GDP ratio was of 7.2% in Italy and
However, it is important to acknowledge the fact, reported above, that the results of these tests with the real interest rate are extremely sensitive to the specification of the model, namely its deterministic components and dynamic structure. This may be due, for example, to the different orders of integration of the variables. At the aggregate level, no effect was found when using the real rate, which confirms the doubts about the power of the tests at the individual country level.

From all these conclusions, several policy implications may be inferred. The first general inference is that although chapter 4 suggests that only Germany satisfies all sustainability conditions, this does not necessarily comprise significant consequences for the interest rate and therefore economic growth. Only in Italy, unsustainability of the central government fiscal policy seems to put pressure on the interest rate, which leads to a further fiscal deterioration, through its effects on the interest expenditures, in a self-fulfilling process.

One second important conclusion from the tests is that in general, apart from the mentioned case of Italy, there seems to be no reason to fear the crowding out effects of deficits, caused by their pressure on the interest rate. Therefore, a policy of fiscal consolidation, especially in an economic recession, would further deteriorate the economic conditions without reducing by itself the interest rate (see, for example Tobin, 1993).

One other meaningful policy implication from this chapter's findings concerns the actions of the monetary authorities. If the central bank maintains a very tight monetary policy, with high interest rates to confront inflationary pressures, this will put a burden on the government's budget accounts. The consequences are harder for those countries with a large stock of public debt. On the one hand, these countries are forced to sell relatively larger amounts of indexed and short-term debt (Missale and Blanchard, 1994), which magnifies the repercussions of monetary policy on the budget. On the other hand, the considerable amount of interest payments on this debt will demand rising primary surpluses or seigniorage revenues in order to keep the global deficit at low levels.

5.2% in Belgium, while all other countries present an average value below 2.3%.

48 De Grauwe (1998) argues that the EU policy mix of monetary and fiscal restriction during the 1990s led to higher debt-GDP ratios, contrary to what happened in the US, where a policy mix of fiscal restriction but monetary ease was able to reduce the deficit and debt.
This compensation for a growing interest burden is certainly more difficult in a monetary union. In the first place, the member states no longer control their seigniorage revenues. Secondly, the primary surplus, usually slow to react, could be even more constrained in a monetary union, due to the tax harmonisation policies, and the tendency for the uniformisation of some spending categories, notably wages and social transfers, education and economic infrastructures (Commission of the EC, 1990: p. 24).

The lack of empirical support for the proclaimed adverse consequences of the deficits on the interest rate suggests the feasibility of a more flexible fiscal policy in most countries, allowing the deficit, for example, to reflect the general economic conditions. However, the evidence equally suggests that, together with this flexibility, the fiscal authorities should also aim at reducing the stock of public debt. With liberalised capital markets, the real interest rate seems to react to the level of debt, with consequences for the economy, in general, and for the budget dynamics, in particular.
7. CONCLUSION

The progressive deterioration of public finances in most industrialised countries has motivated an extensive debate over the long-run feasibility and consequences of persistently undisciplined fiscal policies. The literature is far from conclusive, justifying the importance of complementing the discussion with further empirical evidence. This research area remains particularly relevant given the recent tendency in several countries to impose institutional restrictions on the size and financing of government deficits and debt. These issues have been investigated in this study, with an empirical application to six member-countries of the European Union during the post-war period.

The existing literature on sustainability of fiscal policies has been extended in several directions. First, the dissertation considers a longer time period on a parallel study of several European Union countries. Second, it provides a comparative analysis of central and general government data, highlighting the growing importance of the social security funds. Third, it reconciles the two major strands of the empirical literature on sustainability, allowing the complementary testing of the transversality condition and the collateral constraint. Fourth, this study employs various recent econometric techniques, not applied before in this research area. Finally, it extends the sustainability analysis to assess the effects of undisciplined fiscal policies on monetization, inflation and interest rates.

The first part of the dissertation examines the long-run sustainability of the current path of fiscal policies. The analysis is centred on the intertemporal budget constraint (IBC), an accounting identity requiring the current value of debt to equal the expected discounted sum of all future net-of-interest surpluses or, equivalently, the discounted stock of debt to approach zero as time approaches infinity. Sufficient conditions for sustainability can be derived from this constraint, involving the empirical assessment of univariate or multivariate restrictions on the data generating processes followed by the relevant fiscal variables.
Several alternative methodologies have been proposed to examine this issue, since the seminal paper by Hamilton and Flavin (1986). After surveying the literature on the subject, a specific sequential testing strategy was adopted for the empirical applications. This strategy allows nesting different approaches, evaluating not only the intertemporal constraint but also the stability of the debt-GDP ratio, a supplementary condition for sustainability according to one strand of the literature.

The results of the econometric tests performed in chapter four, with central government data, reveal that only Germany and the Netherlands follow a sustainable fiscal policy. However, in the latter country the complementary condition of a bounded debt ratio is not satisfied, insinuating potential future problems in marketing its debt. For all the other countries, the evidence overwhelmingly indicates unsustainable policies. This suggests the need for a change in policy, to adjust the processes generating either the budget variables or the other relevant macroeconomic time series which set the dynamics of the IBC, such as the growth, inflation and interest rates.

The question of how the budget should be, or has been, balanced intertemporally is closely related to the statistical causality between the government's total expenditures and revenues. The causality analysis reveals that different approaches have predominated in Germany and the Netherlands to achieve intertemporal budget balance. While in the former country fiscal imbalances are usually counteracted by expenditure cuts, in the latter the adjustment burden generally falls on the revenue side. This may help explain the increasing weight of the central government sector in the Dutch, but not the German, economy throughout the post-war period.

For the remaining countries, the causality tests display significant results in only one case, confirming the previous general findings of unsustainable fiscal policies. In Italy, the tests reveal that total revenues Granger-cause expenditures, suggesting that a policy of fiscal consolidation concentrated on raising additional revenues are unlikely to be successful, since it will be at least partially offset by increased spending. The appropriate policy should therefore focus primarily on reforming the expenditure side.

All the tests have been equally performed with two extra data sets. One increased the number of data points available, by using quarterly data. No significant additional information emerged, apart from the confirmation that fiscal variables are highly
vulnerable to nonstationary seasonal behaviour. The other data set changed the focus of analysis to a broader level of public accounts, by considering general government data. In this case, the results also confirm the earlier findings of a sustainable fiscal policy only in Germany and the Netherlands. However, while the latter country now satisfies both sustainability conditions of intertemporal budget balance and bounded debt-GDP ratio, in Germany the second condition is not verified.

This inversion of the results reflects primarily the inclusion of the social security budget in the analysis. Germany's social security deficits, although not endangering the overall sustainability of the general government's fiscal policy, may lead to an ever-increasing debt ratio and the consequent potential problems to place this debt on the market. For the Netherlands, on the other hand, the better results with general government data probably reflect the significant structural reforms undertaken in the social security system since the early eighties.

Interestingly, the non-causality tests also reveal different results with this broader data set. In Germany the burden of adjustment now falls predominantly on the revenue side, possibly explaining why the share of the general government in the economy has been increasing continuously during the whole period. The same happens in France and the Netherlands, denoting the lower elasticity of the social security expenditures.

These empirical tests of sustainability are motivated essentially by the potential economic consequences of a systematic lack of fiscal discipline, one of the most discussed issues among economic academics and policymakers. First, because the government may be tempted to default on its debt, precipitating a financial crisis. Second, due to political pressures upon the central bank to adopt a more accommodative monetary policy, with inflationary consequences. Finally, the standard macroeconomic models suggest that large deficits generally cause high interest rates, with crowding-out consequences on private investment, exports and, consequently, economic growth.

However, no known study of sustainability has previously extended its scope to consider the consequences of (un)sustainable fiscal policies. These are the object of different strands of research. This thesis attempts to bring together these complementary questions within a unified framework and a common data set. After assessing the issue of sustainability of fiscal policies, the following two chapters have investigated the relation
between the fiscal variables and the inflation and interest rates. The empirical analysis employs essentially a vector autoregressive framework, to identify the feedback effects between the variables.

Chapter five examines the effects of the fiscal stance on monetization, or seigniorage, and inflation. Two alternative measures of seigniorage have been computed and successively employed in the econometric tests, but in neither case was it possible to uncover any positive effect of the deficit on monetization. Similar results were found when using the inflation rate instead of the seigniorage revenues, in order to examine also other inflationary, non-monetization, fiscal effects. On the contrary, the results of the tests suggest that, in some cases, the direction of causality is inverted. In most countries higher inflation exerts a prejudicial effect on the deficit, probably due to the increasing cost of debt service. This effect is particularly noticeable in Belgium and Italy, the two most indebted countries in the group.

Chapter six considers the relationship between the interest rate and the fiscal variables, using different approaches in order to explore distinct but interrelated questions, and obtain more robust conclusions. Overall, the evidence shows that large deficits do not necessarily stimulate soaring nominal or real interest rates. Again, the inverse causal effect seems to prevail, suggesting the need to recognise the importance of the fiscal consequences of monetary policy. Higher nominal interest rates, resulting from a tighter monetary policy, worsen the government’s deficit in almost all countries.

Italy is the only country where in general a bidirectional positive relationship was found between these variables, suggesting the existence of a self-fulfilling cycle of rising interest rates and further fiscal deterioration. The same mechanism emerges in Belgium, but only when general government data is considered. These two exceptions coincide, again, with the two most indebted countries in the sample.

If the analysis is performed using the series of debt instead of the deficit, the results are completely different, although less robust. On the one hand, there is a negative correlation between the series of debt and the nominal interest rate, clearly discernible in a more informal graphical analysis. On the other hand, the level of debt has a positive effect on the real interest rate in four of the six countries considered. This suggests that although the financial markets do not seem to react to the level of the deficit, they are
sensitive to the accumulated amounts of debt. The exceptions, France and Italy, have protected their domestic financial markets for a longer period of time, liberalising completely the movements of capital only in the early nineties.

The overall conclusion from a joint assessment of all the above results indicates that fiscal imbalances do not necessarily involve adverse long-run effects on monetization, inflation or interest rates, even if the fiscal situation can be considered unsustainable. This finding contradicts the conventional macroeconomic theory, but is consistent with alternative models of economic behaviour discussed throughout the thesis.

This conclusion suggests the possibility of a higher flexibility of fiscal policies. In the particular case of the European Union, it supports the criticisms on the Maastricht’s deficit convergence criteria and the even stricter rules outlined in the ‘Stability and Growth Pact’. On the one hand, it demonstrates the ineffectiveness of the criteria. Although all countries in the sample have been considered able to join EMU (the UK chose not to do so), only Germany and the Netherlands present sustainable fiscal policies. Instead of concentrating on the size of the deficit and debt at any particular point in time, it is perhaps more informative to examine the IBC. On the other hand, it suggests its undesirability. The criteria were proposed fundamentally as a mechanism to protect the monetary policy. However, the above stated results indicate that, in general, it is the actions of monetary policy which have more significant fiscal consequences.

Nevertheless, the empirical support for higher flexibility does not imply endorsing the continuation of persistently large deficits and debt, for three main reasons. First, current fiscal policies are in general unsustainable in the long-run, requiring a change in the fiscal stance. Second, the expected evolution in demographic structures indicates considerable problems in the future, requiring major fiscal reforms. Third, although the results suggest that deficits do not generally imply adverse effects, the amount of debt does. It restricts the flexibility of fiscal policies by absorbing a considerable fraction of public revenues, may cause a self-fulfilling cycle of deficit and interest rate increases, and influences the level of the real interest rate, jeopardising growth prospects.

Heavily indebted countries are invariably at the mercy of the financial markets. A sudden loss of confidence immediately precipitates rising interest rates and capital flight. The larger the stock of debt, the higher the probability of monetization (or default), and
therefore the higher the long-term real interest rate that will compensate the investors for the expected increase in inflation. The empirical results obtained above confirm this effect. Even if the monetary authorities have not used inflation in the past to ease the burden of deficits and debt, and the empirical results also suggest the non-existence of this effect, that risk nevertheless persists.

In sum, two major tentative policy implications can be derived from these results. One is the desirability for increased coordination between monetary and fiscal authorities, to take into account the long-run mutual interdependencies of their actions. The other is the possibility of allowing more flexible fiscal policies but, at the same time, aiming at low levels of debt. A particular effort seems to be necessary with the reform of the social security system, a growing source of public liabilities in most countries due to the increasingly expensive welfare programs and the unfunded pension schemes.

The empirical analysis in this thesis displays inevitably several weaknesses, suggesting a cautious interpretation of the results and, at the same time, encouraging some potential directions for future research in this area. For example, the results in chapters five and six may be biased due to the omission of potentially important variables in the estimated models. The analysis in these chapters has been performed on a bivariate, occasionally trivariate, vector autoregressive model. The models were kept intentionally parsimonious in the number of variables to avoid the arbitrary choice of other relevant aggregates, or the subjective selection of an underlying macroeconomic model to identify the major determinants in the relation.

Another important factor to consider are the various issues surrounding the choice of interest rate employed in the empirical tests. First, the long-run interest rate on government bonds may not be the most representative rate. Second, the tests should ideally employ the expected inflation rate to compute the \textit{ex ante} real interest rate. The actual values of these variables may not constitute the most adequate proxy. Finally, it may also be argued that an after-tax interest rate should be used, being the rate considered by the private agents when making their decisions. However, the construction of these alternative series faces overwhelming obstacles when a long period is considered.

A potentially more important drawback of the methodology adopted here to test long-run sustainability is that, being based on the past behaviour of the series, it may not take into
account a structural change which might have occurred towards the end of the sample, or is expected to occur in the future. In the particular case of the European Union, as more recent data becomes available, and once sufficient data points can be collected, it would be interesting to examine the influence of EMU on these questions. It is possible that the Maastricht’s fiscal convergence criteria have already provided the necessary adjustment to most unsustainable fiscal policies. However, evidence from the application of formal fiscal limits in the US has shown that they may be ineffective, since governments are often able to circumvent them.

One central element in this question is the behaviour of the interest rates, a crucial variable in the budget dynamics. If the monetary union leads to lower interest rates, and therefore lower interest expenditures, this will improve the fiscal situation. However, if it leads to an increase in the default risk premium, EMU could exacerbate the differences in the fiscal positions of the different member states. Ultimately, the financial markets have a decisive role in the sustainability of fiscal policies.

The main directions for further research should focus primarily on the economic consequences of unsustainable fiscal policies. The tests of the sustainability hypothesis in chapter four have followed extensive and careful methodological procedures, and essentially only new data may provide significant additional information. Two main points may however be worth pursuing. A promising strand of research concerns the application of recently developed panel data cointegration tests to the hypothesis of EU-wide sustainability. There may also be some scope for improvement in the definition and measurement of the fiscal and other economic variables, increasing the quality of the data set, although at the cost of a shorter time period and consequently lower power of the econometric tests.

There is however the need for additional research on the effects of unsustainability. In spite of the voluminous literature examining the fiscal consequences on inflation and interest rates, the evidence is far from conclusive. In particular, future work on this topic could consider the sensitivity of the results in chapters five and six to the inclusion of alternative groups of variables in the models. It would also be very interesting to explore the effects of unsustainability on economic growth, another fundamental variable in the intertemporal budget constraint.
APPENDIX
<table>
<thead>
<tr>
<th>Paper identification</th>
<th>Auxiliary Assumptions</th>
<th>Sustainability Condition</th>
<th>Sample</th>
<th>Conclusion 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton, Flavin 1986</td>
<td>constant, NS stationary</td>
<td>D, NS stationary</td>
<td>C, A</td>
<td>1960-84</td>
</tr>
<tr>
<td>Kremers 1988</td>
<td>constant, NS stationary</td>
<td>D stationary</td>
<td>C, A</td>
<td>1960-84</td>
</tr>
<tr>
<td>Trehan, Walsh 1988</td>
<td>constant, x=(G,T,ΔM/P) - I(1), (1-L)x = μ + ϕ(L)ε,</td>
<td>(rΔ-NS) stationary; D, NS cointegrated [r-1];</td>
<td>C, A</td>
<td>1980-86</td>
</tr>
<tr>
<td>Grilli 1989</td>
<td>constant, G, TT ~ I(1)</td>
<td>rD-NS stationary</td>
<td>C, A</td>
<td>±1948-86</td>
</tr>
<tr>
<td>Wilcox 1989</td>
<td>variable, D ~ general ARIMA</td>
<td>D (discounted) stationary and unconditional mean = 0</td>
<td>C, A</td>
<td>1960-84</td>
</tr>
<tr>
<td>MacDonald, Speight 1990</td>
<td>constant, D, NS both I(0) or I(1)</td>
<td>[D, NS stationary or cointegrated]</td>
<td>G, Q</td>
<td>1961-86</td>
</tr>
<tr>
<td>Bohn 1991</td>
<td>constant, t, g, d ~ I(1)</td>
<td>(rd-s) stationary</td>
<td>C, A</td>
<td>1800-988</td>
</tr>
<tr>
<td>Corsetti 1991</td>
<td>variable, D ~ general ARIMA</td>
<td>D (discounted) stationary or mean=0</td>
<td>P, Q</td>
<td>1975-88</td>
</tr>
<tr>
<td>Grilli et al 1991</td>
<td>constant, g, t ~ I(1)</td>
<td>(rd-s) stationary</td>
<td>C, nd</td>
<td>n.d.</td>
</tr>
<tr>
<td>Hakkio, Rush 1991</td>
<td>stationary with unconditional mean=r, TT, TG ~ I(1)</td>
<td>(G+rD), TT cointegrated</td>
<td>C, Q</td>
<td>1950-88</td>
</tr>
<tr>
<td>Paper identification</td>
<td>Auxiliary Assumptions</td>
<td>Sustainability Condition</td>
<td>Sample</td>
<td>Conclusion¹</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>-----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Haug</td>
<td>constant</td>
<td>( D, NS \sim I(1) ) ( D ) and ( NS ) cointegrated</td>
<td>C Q 1960-87 US</td>
<td>S</td>
</tr>
<tr>
<td>Roberds</td>
<td>constant</td>
<td>Cross-equation restrictions on the VAR representation</td>
<td>C Q 1948-86 US</td>
<td>U</td>
</tr>
<tr>
<td>Smith, Zin</td>
<td>constant</td>
<td>( D, NS ) cointegrated with vector ([1,-1/r]) ( rD-NS ) stationary</td>
<td>C M 1946-84 Canada</td>
<td>U</td>
</tr>
<tr>
<td>Trehan, Walsh</td>
<td>constant</td>
<td>( NS \sim I(1) ) ( D ) and ( NS ) cointegrated</td>
<td>C A 1960-84 US</td>
<td>U</td>
</tr>
<tr>
<td>Buiter, Patel</td>
<td>variable</td>
<td>( d \sim general \ ARIMA ) ( d ) (discounted) stationary and mean=0</td>
<td>P A 1970-87 India</td>
<td>U</td>
</tr>
<tr>
<td>Caporale</td>
<td>variable</td>
<td>( n.d. ) ( d ) (discounted) stationary</td>
<td>n.d Q 1972-91 EU7</td>
<td>U</td>
</tr>
<tr>
<td>Corsetti, Roubini</td>
<td>variable</td>
<td>( D \sim general \ ARIMA ) ( D ) (discounted to base period) stationary or mean =0</td>
<td>P A 1961-89 EU10</td>
<td>U</td>
</tr>
<tr>
<td>Jondeau</td>
<td>constant</td>
<td>( ns \sim (1-\phi L)ns = \theta(L)e_t ) ( g+\rho d ) and t cointegrated with vector ([1,-1]) ( \rho d-ns ) stationary</td>
<td>P Q 1965-90 France</td>
<td>U</td>
</tr>
<tr>
<td>MacDonald</td>
<td>constant</td>
<td>( NS \sim I(0) ) or ( I(1) ) ( D ) and ( NS ) cointegrated or ( rD-NS ) stationary</td>
<td>C M 1951-84 US</td>
<td>U</td>
</tr>
<tr>
<td>Baglioni, Cherubini</td>
<td>constant</td>
<td>( NS \sim I(1) ) ( 1NS ) and ( D ) cointegrated with vector ([1,-r]) ( D ) (discounted) stationary, with mean=0</td>
<td>n.d M 1979-91 Italy</td>
<td>U</td>
</tr>
<tr>
<td>Caporale</td>
<td>constant</td>
<td>( S \sim (1-\phi L)NS = \theta(L)e_t ) ( g ) and ( nt ) cointegrated with vector ([1,-1])</td>
<td>G Q 1971-90 EU10</td>
<td>not clear</td>
</tr>
<tr>
<td>Caporale</td>
<td>variable</td>
<td>( n.d. ) ( D ) (discounted) stationary with mean=0</td>
<td>n.d Q 1972-90 EU7</td>
<td>U</td>
</tr>
<tr>
<td>DeHaan, Siermann</td>
<td>constant</td>
<td>( n.d. ) ( (\rho d-ns) ) stationary</td>
<td>C A 1900-88 Netherl.</td>
<td>S</td>
</tr>
<tr>
<td>Paper identification</td>
<td>Auxiliary Assumptions</td>
<td>Sustainability Condition</td>
<td>Sample</td>
<td>Conclusion¹</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Heinemann</td>
<td>constant ( G, T - I(1) ) ( (G+rD) ) and ( T ) cointegrated with vector ([1 1]) ( \text{US} ) ( F ) ( 1954-88 )</td>
<td>( \text{EU 6} )</td>
<td>( U:B,F,N )</td>
<td></td>
</tr>
<tr>
<td>McNellis, Siddiqui</td>
<td>constant ( \text{variable} ) ( NS - \frac{1}{(1-\phi L)}NS = \theta(L)\varepsilon ) ( (rD-NS) ) and ((1-L)D ) stationary; ( D ) and ( NS ) cointegrated</td>
<td>( n.d ) ( \text{Q} ) ( 1975-n.d. ) ( \text{New Zealand} )</td>
<td>( S )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Wickens</td>
<td>variable ( \text{discounted} NS - I(0) ) or ( I(1) ) ( d ) and ( ns ) (both discounted back to period zero) stationary</td>
<td>( n.d ) ( \text{Q} ) ( 1963-91 ) ( \text{UK} )</td>
<td>( U )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Fukuda, Teruyama</td>
<td>constant ( TG,T - I(1) ) ( (G+rD) ) and ( T ) cointegrated</td>
<td>( C ) ( A ) ( 1888-944 ) ( \text{Japan} )</td>
<td>( U )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Tanner, Liu</td>
<td>stationary with mean = ( r ) ( TG,T - I(1) ) ( (G+rD) ) and ( T ) cointegrated ( (rD - S) ) stationary</td>
<td>( n.d ) ( \text{Q} ) ( 1950-89 ) ( \text{US} )</td>
<td>( S )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Ahmed, Rogers</td>
<td>marginal rate of substitution consumption ( T,G - I(1) ) utility function time-separable; risk premia time-invariant ( T,G ) and ( rD ) cointegrated with vector ([1 -1 -1])</td>
<td>( n.d ) ( \text{A} ) ( 1692-992 ) ( \text{UK} )</td>
<td>( S )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Caporale</td>
<td>constant ( NS - \text{general ARIMA} ) West's (1987) specification test - compares two sets of estimates</td>
<td>( n.d ) ( \text{A} ) ( \pm 1970-91 ) ( \text{EU 10} )</td>
<td>( S )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Haug</td>
<td>stationary with mean = ( r ) ( TG,T,D - I(1) ) ( (G+rD) ) and ( T ) cointegrated with vector ([1 b, 0 &lt; b \leq 1])</td>
<td>( C ) ( Q ) ( 1950-90 ) ( \text{US} )</td>
<td>( U )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Liu, Tanner</td>
<td>n.d. ( TG,T,D - I(1) ) ( (G+rD) ) and ( T ) cointegrated with vector ([1 -1] ) ( G ), ( T ) and ( D_{-1} ) cointegrated with vector ([1 -1 r])</td>
<td>( n.d ) ( \text{Q} ) ( 1956-88? ) ( \text{US} )</td>
<td>( S )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Quintos</td>
<td>stationary around the mean ( TG,T - I(1) ) strong: ( (G+rD) ) and ( T ) cointegrated with vector ([1, b] ), ( 0 &lt; b \leq 1 ) ( weak: just ) ( 0 &lt; b \leq 1 ) even if not cointegrated</td>
<td>( C ) ( Q ) ( 1947-92 ) ( \text{US} )</td>
<td>( S )</td>
<td>( U:B,F,G )</td>
</tr>
<tr>
<td>Paper identification</td>
<td>Auxiliary Assumptions</td>
<td>Sustainability Condition</td>
<td>Sample</td>
<td>Conclusion</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>Roubini</td>
<td>variable</td>
<td>$D \sim \text{general ARIMA}$</td>
<td>$D$ (discounted to base period) stationary or mean=$0$</td>
<td>P</td>
</tr>
<tr>
<td>Fountas, Wu</td>
<td>stationary</td>
<td>$TG,T \sim I(1)$</td>
<td>(G+rD) and $T$ cointegrated</td>
<td>n.d</td>
</tr>
<tr>
<td>Hénin, Garcia</td>
<td>constant</td>
<td>$tt, g, d \sim I(1)$</td>
<td>$tt, g, d$ (pd) cointegrated with vector $[I-1]$; $tt, tg$ cointegrated with vector $[1-1]$</td>
<td>G B</td>
</tr>
<tr>
<td>Uctum, Wickens</td>
<td>variable</td>
<td>discounted $NS \sim I(0)$ or $I(1)$</td>
<td>$d$ (discounted) stationary and mean = $0$</td>
<td>n.d</td>
</tr>
<tr>
<td>Crowder</td>
<td>stationary, constant</td>
<td>$D \sim I(1)$</td>
<td>(G+rD) and $T$ cointegrated with vector $[I-1]$</td>
<td>C Q</td>
</tr>
<tr>
<td>Artis, Marcellino</td>
<td>variable</td>
<td>$d, \delta, d_{t+1} \sim I(1)$</td>
<td>$D, d$ (discounted or not) stationary and mean=$0$; log $\delta$ ($\delta t$) and log $d$ ($\delta t$) cointegrated with vector $[1-1]$</td>
<td>G A</td>
</tr>
<tr>
<td>Bohn</td>
<td>variable</td>
<td>$s$ responds positively to $d; d \sim I(0)$</td>
<td>C A</td>
<td>1916-95</td>
</tr>
<tr>
<td>Makrydakis, Tzavalis, Balfoussias</td>
<td>constant, positive</td>
<td>$d, pd-s \sim I(1)$</td>
<td>$\Delta d$ and $(pd-s)$ stationary with mean=$0$</td>
<td>C A</td>
</tr>
</tbody>
</table>

1 Level of government, central (C), general (G), public (P). 2 Frequency of the observations, annual (A), biannual (B), quarterly (Q) or monthly (M). 3 The countries are identified by their initials, the digits indicate the number of countries included. 4 (US) indicates (Un)Sustainability, while ? indicates inconclusiveness - only the US and the six EU countries are identified here. 'n.d.' indicates 'not defined'.
Figure A.1: Total Revenues and Total Expenditures, Central Government (% GDP)
Figure A.2: Total Surplus, Central Government (% GDP)

Belgium

France

Germany

Italy

Netherlands

United Kingdom
Figure A.3: Total Debt, Central Government (% GDP)
Figure A.4: Total Revenues and Total Expenditures, General Government (% GDP)

Belgium

France

Germany

Italy

Netherlands

United Kingdom

--- Total expenditures

----- Total revenues
Figure A.5: Total Surplus, General Government (% GDP)
Figure A.6: Total expenditures and revenues (% GDP) - quarterly data

Belgium

France

Germany

Italy

Netherlands

United Kingdom

--- Total expenditures

--- Total revenues
Figure A.7: Total expenditures and revenues (% GDP) - quarterly data, seasonally adjust.

Belgium

France

Germany

Italy

Netherlands

United Kingdom

--- Total expenditures ---- Total revenues
Figure A.8: Seigniorage Revenues - Components (% GDP)

Belgium

France

Germany

Italy

Netherlands

United Kingdom

\[ \Delta m_t \]

\[ \frac{\psi_t}{(1 + \psi_t)} m_{t-1} \]

\[ \frac{\pi_t}{(1 + \psi_t)(1 + \pi_t)} m_{t-1} \]
Figure A.9: Seigniorage Revenues - Total (% GDP)

<table>
<thead>
<tr>
<th>Country</th>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
<th>Italy</th>
<th>Netherlands</th>
<th>United Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seigniorage (cash-flow)</td>
<td>![Graph](Belgium cash-flow)</td>
<td>![Graph](France cash-flow)</td>
<td>![Graph](Germany cash-flow)</td>
<td>![Graph](Italy cash-flow)</td>
<td>![Graph](Netherlands cash-flow)</td>
<td>![Graph](United Kingdom cash-flow)</td>
</tr>
<tr>
<td>Seigniorage (opportunity cost)</td>
<td>![Graph](Belgium opportunity cost)</td>
<td>![Graph](France opportunity cost)</td>
<td>![Graph](Germany opportunity cost)</td>
<td>![Graph](Italy opportunity cost)</td>
<td>![Graph](Netherlands opportunity cost)</td>
<td>![Graph](United Kingdom opportunity cost)</td>
</tr>
</tbody>
</table>

Legend:
- **Solid line**: Seigniorage (cash-flow)
- **Dashed line**: Seigniorage (opportunity cost)
Figure A.10: Debt, total surplus and interest rates

Belgium

France

Germany

Legend:

- Debt (right scale)
- Total surplus
- Nominal interest rate
- Real interest rate
Figure A.10: Debt, total surplus and interest rates (cont.)

Italy

Netherlands

United Kingdom

Legend:
- Debt (right scale)
- Total surplus
- Nominal interest rate
- Real interest rate
Figure A.11: Average debt, total surplus and interest rates
REFERENCES


258


*OECD Economic Surveys* (several issues) Organisation for Economic Co-operation and Development.


PERRON, P. (1989a) The great crash, the oil price shock, and the unit root hypothesis, Econometrica, 57(6), 1361-1401.


TANNER, E. and LIU, P. (1994) Is the budget deficit 'too large'?: some further evidence, Economic Inquiry, 32, 511-518.


