REVISTA DE MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA (21). Páginas 103–116.
Junio de 2016. ISSN: 1886-516X. D.L: SE-2927-06.
URL: http://www.upo.es/RevMetCuant/art.php?id=118

Which Countries Pay More or Less for Their Long Term Debt? A CART Approach

González-Fernández, Marcos
Department of Business Economics and Management
Faculty of Economics and Business, University of León (Spain)
E-mail: mgonf@unileon.es

González-Velasco, Carmen
Department of Business Economics and Management
Faculty of Economics and Business, University of León (Spain)
E-mail: mcgonv@unileon.es

ABSTRACT

The objective of this paper is to classify a group of EMU countries according to the main determinants of long-term sovereign bond yields. We apply the Classification and Regression Tree method (CART). According to the findings, countries with lower inflation, a lower debt to GDP ratio, a lower average income tax rate, higher public debt maturity and higher IPI growth are placed in classification groups that have lower bond yields. These results confirm the hypothesis that countries with better macroeconomic and fiscal indicators have lower sovereign bond yields.

Keywords: long-term yields; sovereign yields; classification trees; decision trees.
JEL classification: G12; G15; H63; C38.
MSC2010: 05C05; 62G99; 62H30.

Artículo recibido el 05 de noviembre de 2015 y aceptado el 24 de junio de 2016.

¿Qué países pagan más o menos por su deuda a largo plazo? Una aproximación a través de la metodología CART

RESUMEN

El objetivo de este artículo es clasificar un grupo de países de la UME teniendo en cuenta los principales determinantes de los tipos a largo plazo de la deuda soberana. Se aplica la metodología basada en árboles de decisión. Según los resultados, los grupos de países que tienen menor inflación, deuda pública, tipo impositivo medio y mayor vencimiento de la deuda pública y crecimiento económico pagan menos por su deuda soberana a largo plazo. Se confirma la hipótesis de que los países que tienen los mejores indicadores macroeconómicos y fiscales son los que presentan menores costes en su deuda soberana.

Palabras claves: tipos a largo plazo; rendimientos soberanos; árboles de clasificación; árboles de decisión.
Clasificación JEL: G12; G15; H63; C38.
MSC2010: 05C05; 62G99; 62H30.

1. INTRODUCTION

One of the Maastricht Treaty's convergence criteria that are used for valuing countries that are in the process of entering the European Monetary Union (EMU) is the long-term interest rate. The increased harmonization of monetary and fiscal policies and the adoption of a common currency contributed to the convergence of long-term government bond yields in the EMU. This period of convergence lasted until the collapse of Lehman Brothers in September 2008. The effects of the global financial crisis moved into the real economy, and macroeconomic indicators worsened in many EMU countries, leading to an increase in long-term sovereign bond yields, especially in those countries with high deficits or with a banking sector weakened by the global financial crisis (Greece, Ireland, Portugal, Spain and Italy).

Most studies on advanced economies find empirical support for the theoretical prediction that sovereign debt and other macroeconomic fundamentals have an impact on government bond yields (Caporale and Williams, 2002; Rault and Afonso, 2011; Poghosyan, 2012; Bernoth and Erdogan, 2012).

The main aim of this paper is to classify a group of EMU countries by considering the main determinants of long-term sovereign bond yields to determine which countries pay more or less to borrow in the long run. For this purpose, we apply the Classification and Regression Tree (CART) methodology. This approach allows classifying individuals according to a set of variables of different nature. This methodology, which is commonly applied in other fields such as medicine or biology, is not usually employed in the economic field. Only a few papers deal with this technique. Oral *et al.* (1992) use CART procedure to analyze the determinants of country risk for a set of countries during the 80's. They find evidence that the variables that mainly affect ratings are GDP per capita and the investments to GDP ratio. A more recent example is the analysis of Manasse and Roubini (2009) of sovereign debt crises through CART. They find that high debt and high inflation, along with illiquidity factors driven by large stocks of short debt are the factors that better explain the presence of debt crises. Following this line of research, this paper contributes to the literature with a new perspective in the analysis of sovereign bond determinants applying CART procedure. This technique allows us to predict the value of long term bond yields according to a set of variables and classify the countries in groups with certain confidence intervals according to their expected values. The CART captures nonlinearity in the data and better handle missing data than regression techniques (Morrison, 1998). Thus, economic authorities can know what factors

drive sovereign bond yields, and they can have an estimation of their expected values. In addition, they can predict in which group their country will be whether the conditions or the values of those variables change.

To be exhaustive in the analysis, we consider the most used variables in the literature as influential factors on long-term government bond yields: macroeconomic fundamentals (Hodgson et al., 1998; Hardouvelis, 1998; Kiani, 2009; Gruber and Kamin, 2012), fiscal variables (Ardagna *et al.*, 2007; Laubach, 2009; Maltritz, 2012) and financial indicators (Schuknecht *et al.*, 2009; Afonso et al., 2011; Bernoth and Erdogan, 2012).

2. DATA AND METHODOLOGY

We consider 12 EMU countries: Austria, Belgium, Finland, France, Greece, Germany, Ireland, Italy, Luxembourg, Netherlands, Portugal, and Spain. To classify the countries analyzed[1], we consider the following variables:

- Dependent variable: 10-year government bond yields of each country. The data are obtained from Eurostat.
- Explanatory variables: from the existing literature, we identify a set of variables that may determine long-term sovereign bond yields, aggregated into the following groups: (i) macroeconomic fundamentals (Industrial Production Index (IPI), Consumer Price Index (CPI) inflation, and unemployment rate), (ii) fiscal variables (deficit-to-GDP ratio, public debt-to-GDP ratio, debt growth-to-GDP ratio, private debt-to-GDP ratio, and average income tax rate) and (iii) financial indicators (public debt maturity, and sovereign rating). All data are obtained from Eurostat and Organization for Economic Co-operation and Development (OECD) statistics except the sovereign rating. This variable has been compiled from the score that the three main rating agencies, Fitch, Moody's, and Standard & Poor's, made on the credit quality of each country. We transform this score into a quantitative variable according to Remolona *et al.* (2007), whereby we are able to test its impact on sovereign spreads.

We apply the CART methodology, a computer-intensive data-mining technique that selects explanatory variables, their critical values, and their interactions to classify different countries according to the main determinants of long-term sovereign bond yields. This

[1] The data considered are the average values of the variables in the period 2000-2010. For this reason the countries that joined the EMU since 2007 (Slovenia, Cyprus, Malta, Slovakia, Estonia) are not included in the analysis.

technique was developed by Breiman *et al.* (1984). The CART methodology's main field of application is the experimental sciences, especially medicine. In the economic sphere, CART is a more recent application method (Manasse and Roubini, 2009).

This method uses a binary and recursive procedure, whereby parent nodes are split into two child nodes using splitting rules based on predictor variables; the process is repeated to reduce the conditional variation in the response variable. The CART method is used to search for the characteristics that are most closely associated with group membership. The key elements of a CART analysis are a set of rules for

1. Splitting a parent node into two child nodes with questions that have a "yes" or "no" answer. For example, if we use CPI inflation as an explanatory variable, the question could be: "is X country´s CPI inflation higher than 2 percentage points?" The CART method analyzes all possible splits for all included variables and selects the one that best separates the dependent variable; in our case, this variable is long-term sovereign bond yields. In practice, the CART obtains two groups according to the explanatory variables, and the process is repeated within these sub-groups. The CART method calculates an error, and selects the split that minimizes the error with the Gini criterion.
2. Deciding when to stop growing the tree when the reduction in the misclassification rate falls below the penalty associated whether we obtain more nodes.
3. Assigning each terminal node to a group. In our case, the CART algorithm creates groups based on the level of yields in each country and assigns each country to one of the groups.

3. RESULTS

To perform the analysis, we first select long-term sovereign bond yields as the dependent variable[2], and we select the following explanatory variables: unemployment, deficit, public debt-to-GDP ratio, debt growth-to-GDP ratio, CPI inflation, private debt-to-GDP ratio, average income tax rate, IPI, public debt maturity and the rating (Table 1).

[2] We employ the XLSTAT statistical software from Microsoft Excel.

Table 1. Descriptive statistics of variables

Variable	Observations	Minimum	Maximum	Mean	Standard Deviation
IPI	11	-1.327	4.460	1.241	2.062
CPI inflation	11	1.683	2.908	2.277	0.391
Unemployment rate	11	3.885	12.785	7.428	2.647
Deficit-to-GDP	11	-4.650	2.867	-1.666	2.261
Public Debt-to-GDP	11	9.485	108.977	61.562	26.665
Private debt-to-GDP	11	100.425	315.960	178.297	59.747
Public debt growth-to-GDP	11	-0.012	0.120	0.034	0.042
Average income tax rate	11	29.623	44.288	38.593	5.294
Public debt maturity	11	4.136	7.133	5.817	0.959
Rating	11	1.000	5.833	2.240	1.709

Source: own elaboration
All variables are expressed in percentage except the public debt maturity which is in years, and the rating.

The CART selects the following five variables out of the 10 countries listed in Table 2: CPI inflation, public debt/GDP, average income tax rate, IPI, and public debt maturity.

Table 2. CART analysis

Node	Countries	No. of observations	Parent nodes	Split variable	Values
1	All	11 (100%)			
2	AT, BE, FI, FR, IT, NT, GE	7 (63.64%)	1	CPI inflation	[1.683;2.371)
3	SP, IR, LUX, POR	4 (36.36%)	1	CPI inflation	[2.371;2.908)
4	AT, FI, FR, NT, GE	5 (45.45%)	2	Public debt/GDP	[42.569;82.508)
5	BE,IT	2 (18.18%	2	Public debt/GDP	[82.508;108.977)
6	NT, GE	2 (18.18%)	4	Average income tax rate	[36.184; 40.667)
7	AT, FI, FR	3 (27.27%)	4	Average income tax rate	[40.667; 44.04)
8	NT	1 (9.09%)	6	IPI	[1.222; 1.774)
9	GE	1 (9.09%)	6	IPI	[1.774;2.326)
10	FI, FR	2 (18.18%	7	IPI	[-0.244; 2.142)

Node	Countries	No. of observations	Parent nodes	Split variable	Values
11	AT	1 (9.09%)	7	IPI	[2.142; 3.621)
12	FI	1 (9.09%)	10	Public debt maturity	[4.136; 5.288)
13	FR	1 (9.90%)	10	Public debt maturity	[5.288; 6.439)
14	IT	1 (9.09%)	5	IPI	[-0.873; 1.495)
15	BE	1 (9.09%)	5	IPI	[1.495; 3.863)
16	IR, POR, LUX	3 (27.27%)	3	Public debt maturity	[4.285; 5.84)
17	SP	1 (9.09%)	3	Public debt maturity	[5.84; 6.307)

Source: own elaboration
Countries: AT (Austria) ,BE (Belgium), FI (Finland), FR (France), Germany (GE), Ireland (IR), Italy (IT), Luxembourg (LUX), The Netherlands (NT), Portugal (POR), and Spain (SP).

The first rule splits the sample into two child nodes when we use the variable "CPI inflation": (i) node 2 with seven countries (63.64%) with low inflation between 1.683% and 2.371%, and (ii) node three with four countries (36.36%) with high inflation greater than 2.371% (Table 2). Countries with low inflation (node 2) are further split into two other nodes with low/high public debt/GDP ratio: (i) node 4, with five countries (45.45%) with a low public debt/GDP ratio between 42.57% and 82.51%, and (ii) node five, with two countries (18.18%) with a high private debt/GDP ratio greater than 82.51%. Additionally, countries with high inflation (node 3) are split into two terminals when we use the variable "public debt maturity": (i) node sixteen, with three countries (27.27%) with a low public debt maturity between 4.28 and 5.84 years, and (ii) node seventeen, with one country (9.09%) with a high public debt maturity of more than 5.84 years. We repeat this process for the subsequent nodes using the following split rules (Table 3).

Table 3. Split rules into nodes

Node	Predicted yields	Split rules and critical thresholds
1	4.353	
2	4.213	If CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.213% in 63.6% of cases.
3	4.597	If CPI inflation is between 2.371 and 2.908, then the long-term yield is 4.597% in 36.4% of cases.

Node	Predicted yields	Split rules and critical thresholds
4	4.114	If the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.114% in 45.5% of cases.
5	4.460	If the public debt/GDP ratio is between 82.508 and 108.977, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.460% in 18.2% of cases.
6	4.029	If the average income tax rate is between 36.184 and 40.667, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.029% in 18.2% of cases.
7	4.171	If the average income tax rate is between 40.667 and 44.04, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.171% in 27.3% of cases.
8	4.111	If the IPI is between 1.222 and 1.774, the average income tax rate is between 36.184 and 40.667, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.111% in 9.1% of cases.
9	3.948	If the IPI is between 1.774 and 2.326, the average income tax rate is between 36.184 and 40.667, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 3.948% in 9.1% of cases.
10	4.146	If the IPI is between -0.244 and 2.142, the average income tax rate is between 40.667 and 44.04, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.146% in 18.2% of cases.
11	4.222	If the IPI is between 2.142 and 3.621, the average income tax rate is between 40.667 and 44.04, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.222% in 9.1% of cases.
12	4.148	If the public debt maturity is between 4.136 and 5.288, the IPI is between -0.244 and 2.142, the average income tax rate is between 40.667 and 44.04, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.148% at 9.1% of cases.
13	4.143	If the debt maturity is between 5.288 and 6.439, the IPI is between -0.244 and 2.142, the average income tax rate is between 40.667 and 44.04, the public debt/GDP ratio is between 42.569 and 82.508, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.143% at 9.1% of cases.
14	4.584	If the IPI is between -0.873 and 1.495, the public debt/GDP ratio is between 82.508 and 108.977, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.584% in 9.1% of cases.

Node	Predicted yields	Split rules and critical thresholds
15	4.336	If the IPI is between 1.495 and 3.863, the public debt/GDP ratio is between 82.508 and 108.977, and CPI inflation is between 1.683 and 2.371, then the long-term yield is 4.336% in 9.1% of cases.
16	4.640	If the public debt maturity is between 4.285 and 5.84, and CPI inflation is between 2.371 and 2.908, then the long-term yield is 4.640% in 27.3% of cases.
17	4.468	If the public debt maturity is between 5.84 and 6.307, and CPI inflation is between 2.371 and 2.908, then the long-term yield is 4.468% in 9.1% of cases.

Source: own elaboration
All variables are expressed in percentage except the public debt maturity which is in years, and the rating.

The CART method assigns a yield to every node. Therefore, node 9 presents the lowest estimated yields. This node includes the following characteristics: low inflation, reduced public debt/GDP ratio, low average income tax rate, and high IPI. According to this methodology, this combination of variables is most likely to result in a sovereign bond with lower yields. Germany is observed in this node (Table 2), which is unsurprising because it is the reference country to measure the sovereign risk, and it has the lowest default probability. Next is node 8, which includes the Netherlands, which has the same characteristics as Germany but with a lower IPI. In third place is node 13, which contains France.

On the opposite side is node 16, which contains Portugal, Ireland and Luxembourg; this node is characterized by higher inflation and a lower public debt maturity than the others. These characteristics are more likely to indicate a country that has higher yields and is therefore in danger of greater default. Next is node 14, which includes Italy and is characterized by a high debt/GDP ratio and a reduced IPI, despite moderate inflation. In third place is node 17, which contains Spain, a country with high inflation and longer public debt maturity.

Thus, countries with higher inflation, a higher debt/GDP ratio, a higher average income tax rate, lower public debt maturity, and lower IPI growth suffer greater pressure because their chances of having a high yield are superior to those countries in which the opposite conditions are present. This result confirms the hypothesis that countries with higher values in the macroeconomic and fiscal variables (except for IPI and public debt maturity, which present an inverse relationship) have higher yields. This methodology also allows us to determine the quality of the classification tree (Table 4).

Table 4. Quality of the fit of CART

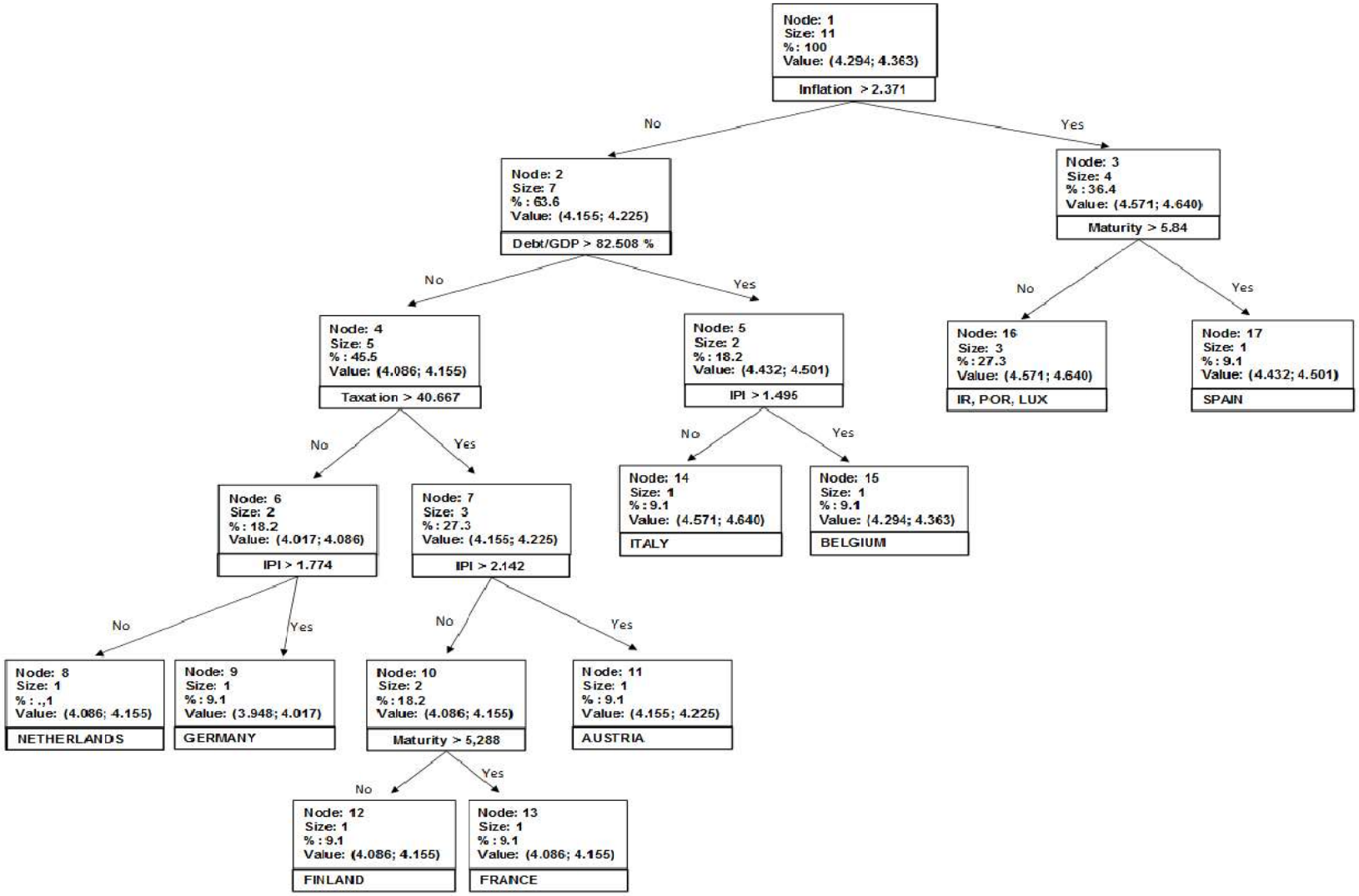
Countries	Yields	Predicted yields
Austria	4.222	4.222
Belgium	4.336	4.336
Finland	4.148	4.148
France	4.143	4.143
Germany	3.948	3.948
Ireland	4.996	4.640
Italy	4.584	4.584
Luxembourg	3.923	4.640
The Netherlands	4.111	4.111
Portugal	5.000	4.640
Spain	4.468	4.468

Source: own elaboration

The CART does not consider one of the countries from the sample and runs the algorithm to obtain the best classification tree. In this case, the CART does not include Greece. Most likely, this finding is due to the special nature of the crisis in Greece; therefore, its inclusion would significantly alter the results of the CART.

The CART adequately classifies 70% of the cases. However, it is noteworthy that the CART correctly classifies 100% of the terminal nodes, except for node 16, which includes Ireland, Portugal and Luxembourg. In this node, Luxembourg is assigned a yield of 4.64 when the actual yield is 3.923, which is lower even than that of Germany; therefore, Luxembourg should be classified in a different terminal node. Additionally, Ireland and Portugal are also assigned lower than their actual yields; these countries should also be included in a worse classification group. Therefore, node 16 seems to misclassify the countries included in it. One possible explanation is that Portugal and Ireland were affected by sovereign debt crisis that heavily stressed their sovereign risk levels. The CART approach classifies them in the node with higher yields, which is correct, due to their financial situation, but it provides lower values of estimated yields than the real ones. It is due to the inclusion of Luxembourg in the same node. Luxembourg is an atypical case with very especial economic and financial characteristics and its inclusion in the node 16 biases the results because it underestimates the yields of this node.

Figure 1. Classification and regression tree



Source: own elaboration

The regression tree (Figure 1) is the basic element of the CART method; the tree presents the different groups of EMU countries according to the main determinants of long-term government bond yields. The node with the lowest yield includes Germany, whereas Luxembourg, Ireland and Portugal are included in the node with higher yields. Each node displays the following information: the values of the splitting variable, the node number, the number of countries, and the percentage of total countries located in the node. A more detailed analysis is obtained by considering Table 3, which summarizes the split rules that generated the CART algorithm. This analysis shows that we start from a node with eleven individual countries with an average yield between 4.29 and 4.36. From this initial node, the CART performs a first segmentation with inflation as the split variable, and we obtain two nodes: the one on the left comprises the better-off countries (countries where the yields are lower than in the initial node), and the node on the right contains the worse-positioned countries (those with yields higher than in the initial node). After that first division, new partitions are made based on other splitting variables, and the process continues until the CART places the analyzed countries into the terminal nodes. Individual countries situated on the left of the initial node are those with lower yields, while the worse-positioned countries are on the right.

4. CONCLUSIONS

In this paper, we classify a group of EMU countries by considering the main determinants of long-term sovereign bond yields through a CART approach. With respect to explanatory power, the CART procedure allows us to classify correctly almost the 75% of the cases. These results are slightly lower than those obtained by Oral *et al.* (1992) around 80%, because we analyze a different period with more uncertainty due to the recent events in sovereign debt markets.

Regarding our findings, those countries with lower inflation, a lower debt/GDP ratio, a lower average income tax rate, higher public debt maturity and higher IPI growth are placed in classification groups that have lower bond yields. The country that comes closest to these ideal conditions is Germany; this result confirms that it is the reference country for the Eurozone followed by the Netherlands. Conversely, countries with higher inflation, a higher debt/GDP ratio, a higher average income tax rate, lower public debt maturity, and lower IPI growth suffer greater pressure because their chances of having a high bond yield are superior to those countries where the opposite conditions are present. The findings are close to those

obtained by Manasse and Roubini (2009). It seems that high indebted countries, with high level of inflation and low growth are more prone to increase sovereign bond yields and get involved in a sovereign debt crisis. However, these outcomes are slightly different to Oral et al. (1992) that obtain evidence that GDP per capita and the investment to GDP ratio are the variables that mostly affect country risk ratings in the 80's. These results confirm the hypothesis that countries with better macroeconomic and fiscal indicators have lower government bond yields.

REFERENCES

Rault, C. and Afonso, A. (2011). "Long-run determinants of sovereign yields", *Economic Bulletin*, 31(1); 367-374.

Afonso, A.; Gomes, P. and Rother, P. (2011). "Short-and long-run determinants of sovereign credit ratings", *International Journal of Finance and Economics*, 16(1); 15.

Ardagna, S.; Caselli, F. and Lane, T. (2007). "Fiscal discipline and the cost of public debt service: Some estimates for OECD countries", *The B.E. Journal of Macroeconomics*, 7(1); 1-28.

Bernoth, K. and Erdogan, B. (2012). "Sovereign bond yield spreads: A time-varying coefficient approach", *Journal of International Money and Finance*, 31(3); 639-656.

Breiman, L.; Friedman, J.; Olshen, R. and Stone, C. (1984). *Classification and Regression Trees*. Chapman & Hall.

Caporale, G.M. and Williams, G. (2002). "Long-term nominal interest rates and domestic fundamentals", *Review of Financial Economics*, 11(2); 119-130.

Gruber, J. W. and Kamin, S.B. (2012). "Fiscal positions and government bond yields in OECD countries", *Journal of Money, Credit and Banking*, 44(8); 1563-1587.

Hardouvelis, G.A. (1988). "Economic news, exchange rates and internal rates", *Journal of International Money and Finance*, 7(1); 23-35.

Hodgson, A.; Kremmer, M.L. and Shane, L. (1998). "Endogeneous Sovereign bond yield spreads: A time-varying coefficient approach", *Journal of International Money and Finance*, 31(3); 639-656.

Kiani, K.M. (2009). “Federal budget deficits and long-term interest rates in USA”, *The Quarterly Review of Economics and Finance*, 49(1); 74-84.

Laubach, T. (2009). “New evidence on the interest rate effects of budget deficits and debt”, *Journal of the European Economic Association*, 7(4); 858-885.

Maltritz, D. (2012). “Determinants of sovereign yield spreads in the Eurozone: A Bayesian approach”, *Journal of International Money and Finance*, 31(3); 657-672.

Manasse, P. and Roubini, N. (2009). “Rules of "thumb" for sovereign debt crises”, *Journal of International Economics*, 78(2); 192-205.

Morrison, J. (1998). “Introducing CART to the forecasting process”, *The Journal of Business Forecasting Methods & Systems*, 17(1); 9-12.

Oral, M.; Kettani, O.; Cosset, J.C. and Daouas, M. (1992). “An estimation model for country risk rating”, *International Journal of Forecasting*, 8(4); 583-593.

Poghosyan, T. (2012). “Long-run and short-run determinants of sovereign bond yields in advanced economies”, IMF Working Papers, 271.

Remolona, E.; Scatigna, M. and Wu, E. (2007). “Interpreting sovereign spreads”, *BIS Quarterly Review*, March; 27-39.

Schuknecht, L.; Von Hagen, J. and Wolswijk, G. (2009). “Government risk premiums in the bond market: EMU and Canada”, *European Journal of Political Economy*, 25(3); 371-384.