Financial distress in European vineyards and olive groves

Mário Santiago Céu*, Raquel Medeiros Gaspar**

DOI: 10.30682/nm2401c
JEL codes: G32, G33, Q14

Abstract

This study focuses on the prediction of financial distress of agricultural firms operating in the vineyards and olive crops sectors in Mediterranean countries, specifically in Portugal, Spain, and Italy, which are crucial for the production of these crops. The sample size of the study is 5,057 firms. Twelve models are presented, estimated from subsamples of combinations between countries and crops. Logistic regression is used for the estimation of these models. The accuracy of the models is evaluated, considering the importance of misclassification costs. Additionally, the areas under the ROC curves are calculated and compared in a dynamic of possible combinations between crops and countries. The study concludes that there are differences between the two sectors, as well as across countries, and suggests that dedicated models for each country or crop may improve the the models' accuracy.

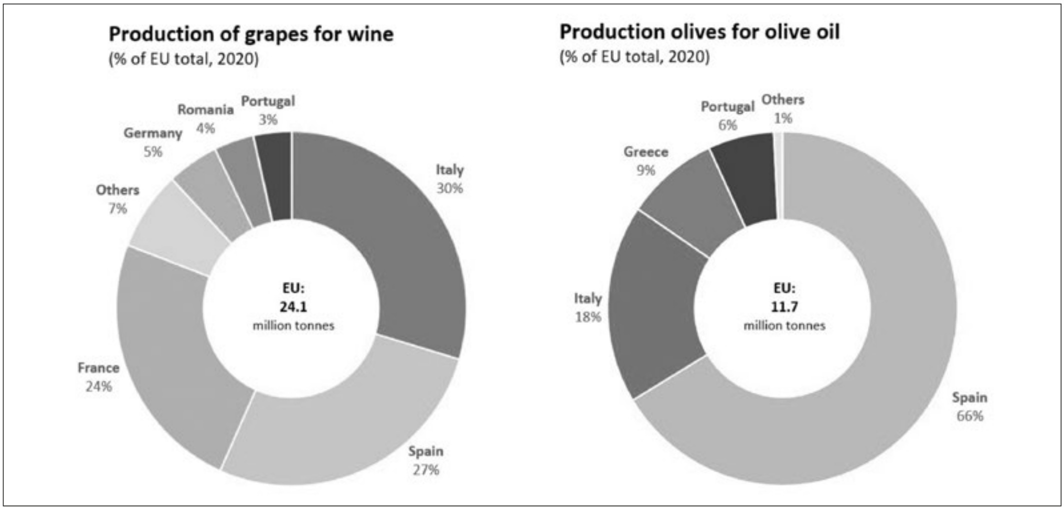
Keywords: *Agriculture, Financial distress, Prediction models, ROC curves.*

1. Introduction

The similarities between the Mediterranean regions in biophysical, climatic and structural conditions are widely recognised. From this similarity, agronomic practices also evolved, predominantly for certain plantations, namely the cultivation of vineyards and olive groves (Caraveli, 2000). "In the Mediterranean basin, the olive along with the vine constituted the equivalent of the rural industries of the North. This equivalence is important, if not for the volume of income, at least for the number of people they engaged, since the 16th century and on, whenever an increase of the cultivation of the olive is observed" (Loumou and Giourga, 2003, p. 90). In 2020, the European Union (EU) explored 3.2 million hectares of vineyards and 5.1 million hectares of olive groves, corresponding to 45% of the world's wine-growing area and 40% of the olive-growing area. From 1962, when the first common market organisation was created, until 2013, when the last reform was revised, the wine sector became more competitive, with simpler and more balanced market rules. European policies over this half century have significantly transformed the sector through diversified interventionist measures, initially supporting divestments (grubbing up) and then supporting firms in financing the restructuring of most of the current vineyards. In 2014, the eight largest EU wine-producing countries accounted for 94% of the EU's wine exports and 65% of global wine exports (Correia *et al.*, 2019). Concerning the production of olives, mainly destined for the extraction of olive oil, the Mediterranean coun-

* ISEG, Lisbon University, Advance/CSG Research Center, Lisboa, Portugal.
** ISEG, Lisbon University, Cemapre/REM Research Center, Lisboa, Portugal.
Corresponding author: mariosantiago@phd.iseg.ulisboa.pt

Figure 1 - Production of grapes and olives in EU.



Source: Eurostat.

tries have had almost absolute dominance in the world due to their unique and highly favourable climate for this culture. In the case of the European Union, Spain, Italy, Greece, and Portugal are the major producers in this market, with a proportion of 99% of the EU-27[1] and 40% of the world, respectively (FAO, 2022). Through Figure 1, it is possible to verify the largest producers of grapes and olives in the European Union.

Despite the importance of these crops in European agriculture and the economy, there are uncertainties about their future. Due to climate change, Fraga *et al.* (2019) refer to risks to the economic sustainability of vineyards and olive groves in these countries. Furthermore, within the olive sector, there is a coexistence of modern and traditional farms, exhibiting significant disparities in productivity, management practices, economic performance, contributions, and sustainable values, raising concerns about the adaptability and survival prospects of traditional family farms (Mokrani *et al.*, 2022). The prevalence of small-scale agriculture also impacts firm viability, wherein farm size and distribution are intrinsically linked to efficiency, with larger farms demonstrating greater productivity and technological advantage, enhancing their survival prospects (Ruz-Carmona *et al.*, 2023).

On the other hand, the Mediterranean countries, compared to Northern Europe, suffer from an ageing agricultural population and poor farm training, which negatively impacts their financial performance. The reformulation of direct payments under the CAP (Common Agricultural Policy), added to the impact of climate change and the liberalization of agricultural trade, places these rural economies in the South more exposed to financial risks (Giannakis and Bruggeman, 2015).

The similarities in the geography and agronomic practices of Portugal, Spain and Italy are widely studied (Arnalte-Alegre and Ortiz-Miranda, 2013; Beopoulos, 2017). However, this does not mean that we can consider a single financial distress prediction model for agricultural firms from different countries and crops.

The paper aims to examine the financial sustainability and risk of agricultural firms, particularly vineyards and olive groves of Portugal, Spain, and Italy. These crops play a vital role in the region's agriculture and economy. However, uncertainties and challenges threaten their future, including climate change risks, disparities between modern and traditional farms, and the impact of policy changes on financial performance. While studying similarities in geog-

[1] The 27 European Union countries after the UK left the EU.

raphy and agronomic practices, a one-size-fits-all approach may not be suitable due to unique economic, social, and environmental factors influencing the financial health of these farms. Hence, are presented financial distress prediction tools for each of these dimensions.

The remainder of the paper is organised as follows. Section 2 addresses the literature on the definition of financial distress, particularly in agriculture, and the relevance of the ROC (receiver operating characteristic) curve to measuring the accuracy of predictive models. Section 3 describes the data and methodology, and section 4 presents and discusses the results. Finally, Section 5 presents the conclusions and limitations.

2. Literature review

After the seminal study by Beaver (1966), the prediction of bankruptcy and financial distress has been a subject of significant interest and research among scholars. While bankruptcy is a legal action that decrees the end of business activity, financial distress results from financial difficulties compromising the firm's ability to honour its commitments. We can define financial distress as a stage before a court decrees bankruptcy. Fitzpatrick (1934) characterizes five moments in the life of a company until bankruptcy: (i) incubation, (ii) embarrassment, (iii) financial insolvency, (iv) total insolvency, and (v) confirmed insolvency. Altman *et al*. (2019) goes deeper into the different concepts and list six reasons that alone or together can contribute to corporate failure, namely (i) poor operating performance and high financial leverage, (ii) lack of technological innovation, (iii) liquidity and funding shock, (iv) relatively high new business formation rates in specific periods, (v) deregulation of key industries, and (vi) unexpected liabilities. The duration between a firm showing signs of financial distress and its bankruptcy being declared is imprecise. However, the years before this failure show predictors of this failure. Chan and Rotenberg (1988) estimated this duration at four years in the Canadian agricultural sector. However, financial distress does not necessarily imply bankruptcy, and many firms prosper after going through moments of financial difficulty.

In the credit risk literature, there are different approaches to defining financial distress, as if it were a singular state dependent on numerous internal or external variables, in addition to different interactions with the local policies and economies in which they operate. "A firm is in financial distress at a given point in time when the liquid assets of the firm are not sufficient to meet the current requirements of its hard contracts" (Hotchkiss *et al.*, 2008, p. 6). Wruck (1990) defines financial distress as an insufficient cash flow to cover current obligations. Asquith *et al.* (1994) bases the entire definition on interest coverage ratios, classifying the firm in financial distress if, for two consecutive years, EBITDA (earnings before interest, taxes, depreciation and amortization) is less than interest expenses or if in one year, EBITDA is less than 80 per cent of its interest expenses. Whitaker (1999) reports this state for the first year in which cash flow is less than current long-term debt maturities. However, one thing is for sure, "distinguishing between financially distressed and healthy companies is more difficult than the traditional comparison between bankrupt and healthy companies" (Platt and Platt, 2006, p. 155).

There are characteristics of the markets and sectors of activity in which firms operate which can compromise the effectiveness of insolvency prediction models. Research on these differences is well known and focuses on various aspects such as cultural, legal, regulatory or macroeconomic. The financial health of firms must be examined in loco within the local macro environment (Khoja *et al.*, 2019). Nevertheless, within similar markets, depending on the sector of activity, there may be variables that stand out as affecting the financial health of firms. In the European Union (EU-27), public policies are shared in the agricultural sector, and even in countries that share similar climates and favourable conditions for the exploitation of certain agricultural products, this does not mean that firms in these countries have similar levels of financial distress.

The lack of a formal definition of financial distress, unlike bankruptcy, which the court defines on a specific date, motivated researchers to propose concepts that somehow characterise the financial strength of firms but emphasise the subjectivity about the most appropriate variables for the definition of this state of the financial health of firms. In the repository of research on financial distress in agriculture, the transnational specificities or the agricultural products cultivated are only sometimes analysed. The data is collected across territories without any differentiation. Klepac and Hampel (2017) tested 250 agriculture business firms in the EU (forestry and logging, fishing and aquaculture), of which 62 reported the default of payment or insolvency proceedings. Vavřina *et al.* (2013) were concerned with homogenizing the data, limiting the choice of 2,581 active and 71 bankrupted agribusiness firms in the Visegrad Group countries (Czechia, Hungary, Poland and Slovakia). Other studies selected firms from agricultural subsectors without proper homogenization criteria. Karas *et al.* (2017) selected 450 active and 25 bankrupt firms. Data were obtained from cereals, rice, grapes, plant propagation, raising of sheep and goats, and mixed farming subsectors. In this selection, they mixed small samples from such subsectors as non-perennial crops, perennial crops and livestock.

Literary approaches that compare predictive models of bankruptcy or financial difficulties in agriculture across various countries and crop combinations are lacking. There is a great diversity of agronomic practices that influence the business structures themselves. The risk of failure for a farmer who produces olives may differ from another farmer who explores vineyards. The same is valid for many other combinations. This research opens a reflection on the subject and aims to contribute to filling this gap.

3. Data and methodology

3.1. *Data and definition of financial distress*

The financial data used in this study is sourced from the Orbis database, provided by Bureau van Dijk. This database is a reputable and widely utilized financial resource, consolidating information from diverse sources, including company reports, regulatory filings, and other publicly available records. It offers extensive financial data for a vast number of companies worldwide. Within the scope of our research, we employed this database to collect financial information about firms operating in the viticulture and oliviculture sectors across European countries. Are considered only firms that did not fail to submit accounts in 2018, 2019 and 2020. We excluded firms that did not have known operating revenue (turnover) in these three years. Of the European countries dedicated to viticulture and oliviculture, only Italy, Spain and Portugal had sufficient financial data available.[2] Table 1 presents the distribution of firms according to the above classifications. We divided data into two groups: just 2018 and both 2019 and 2020. Following the same procedure devised by Platt and Platt (2008), we implemented a two-step procedure to categorize firms according to their financial health. To belong to the healthy group, firms had to register three positive variables in 2019 and 2020. If any of these metrics failed, they would be placed in the financially distressed group; otherwise are categorized as healthy. The variables chosen were (i) EBITDA to interest coverage, (ii) EBIT (earnings before interests and taxes), and (iii) Net income before special items[3]. The financial ratios used to estimate the models are obtained from the 2018 financial statements. This methodology allows us to retrospectively define the status of firms, knowing their performance in the following two years.

Table 2 presents descriptive statistics of applying the two-step procedure to the variables that define the categorization of firms between healthy and financially distressed.

3.2. *Method and hypotheses*

Although the methodology of discriminant analysis gained popularity with Altman (1968), it was from the 1980s onwards that logistic re-

[2] Although there were 399 growing grapes French firms, there were only two firms in olive cultivation. In Greece, only four growing grapes firms were available.

[3] To calculate this last variable and consider the lack of uniformity between the accounting standards, we adopted the formula of extracting extraordinary items (revenue and expenses) from net income.

Table 1 - Distribution of financial statements.

	Vineyards			*Olive Groves*			*Totals*		
	H	FD	% FD	H	FD	% FD	H	FD	% FD
Portugal	738	117	13.7%	351	51	12.7%	1089	168	13.4%
Spain	426	87	17.0%	471	83	15.0%	897	170	15.9%
Italy	1456	399	21.5%	643	235	26.7%	2099	634	23.2%
Totals	2620	603	18.7%	1465	369	20.1%	4085	972	19.2%

Healthy (H), Financial Distressed (FD), (% FD) Proportion of distressed.
Source: Own elaboration.

Table 2 - Descriptive statistics of the firms categorization procedure.

		EBITDA interest coverage[a] 2019		*EBITDA interest coverage 2020*		*EBIT 2019*		*EBIT 2020*		*Net income before special items[b] 2019*		*Net income before special items 2020*	
		Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
Italy													
Vineyards	FD	-141.92	-60.52	-144.30	-52.19	-216.33	-97.97	-259.77	-95.47	-232.76	-101.58	-263.18	-99.96
	H	260.08	31.70	226.62	21.74	130.41	14.30	92.02	8.06	82.91	3.34	62.69	1.63
Olive grows	FD	-52.73	-16.29	-55.14	-15.86	-82.54	-27.03	-88.73	-27.20	-86.59	-28.66	-89.93	-29.60
	H	31.97	4.41	28.65	3.54	-3.21	1.83	9.40	2.67	10.61	0.45	7.18	0.66
Portugal													
Vineyards	FD	-72.19	-19.13	-70.23	-20.89	-86.52	-27.78	-89.81	-30.40	-104.21	-31.50	-107.81	-31.56
	H	94.85	22.93	81.86	17.97	55.07	9.29	38.15	4.93	35.84	5.74	23.09	2.95
Olive grows	FD	-121.49	-51.84	-177.84	-39.05	-149.69	-102.96	-225.64	-62.55	-177.67	-111.17	-271.99	-64.29
	H	119.01	29.78	114.26	30.63	65.64	12.64	67.43	10.98	23.47	8.54	47.32	7.93
Spain													
Vineyards	FD	-61.54	-28.22	-57.07	-28.99	-94.27	-38.54	-87.15	-43.20	-80.69	-37.36	-74.12	-34.00
	H	125.57	29.35	129.40	19.20	82.78	14.71	86.99	6.91	68.92	8.51	65.37	3.70
Olive grows	FD	-71.36	-28.37	-73.77	-20.01	-85.83	-45.76	-88.40	-34.22	-74.47	-36.71	-79.13	-31.40
	H	88.46	29.52	61.67	20.51	59.51	15.77	31.93	10.68	38.85	9.62	59.20	5.84

Healthy (H), Financial Distressed (FD). Source: Own elaboration.
[a]EBITDA interest coverage = EBITDA - Financial expenses. [b]Net income before special items = Net income + Extraordinary and other expenses - Extraordinary and other revenue.

gression came to be preferred by researchers and is even used in the overwhelming majority of bank scorecards (Nyitrai and Virág, 2019). Ohlson (1980) was at the origin of this popularity with his seminal work in the literature on credit risk. In this study, given the characteristics of the sample, namely the disproportion between healthy firms and firms in financial distress, we use binary logistic regression. In logistic regression or a probit model, the model's predictive capacity also depends on definining a cutoff to separate healthy firms from the rest. There is no single way to determine the optimal cutoff. Ohlson (1980) states that previous prediction studies have two assumptions present. First is the presentation of a (mis)classification matrix. Second, an additive property in which the best cutoff point is the one that minimizes the sum of type I (classify a distressed firm as healthy) and type II (classify a healthy firm as distressed) percentage errors. However, it must be considered that comparing models in different periods, predictors, and data sets is exceptionally difficult. Also, the costs are not equal. The cost of classifying a distressed firm as healthy implies losing the return on investment, and the cost of classifying a healthy firm as distressed means losing the investment opportunity (Agarwal and Taffler, 2008). Other authors have tried other approaches. Hsieh (1993) defines type I error as the opportunity cost of holding a long position in equity securities of failing firms. In turn, the type II error is defined as the opportunity cost of selling short securities of healthy firms. Aware of this importance Dopuch *et al.* (1987) and Koh (1992), studied the misclassification costs of type I and type II Errors through proportions from 1:1 to 20:1 and 1:1 to 500:1, respectively. The analysis of type I and type II Errors is very present in the literature. It is indispensable in this kind of research, having the great advantage of being easy to interpret, even for those who do not have a high level of mathematics and statistics education (Čámská *et al.*, 2016). In short, the accuracy of a model goes far beyond the simple calculation of the correct percentage of observation classifications. Moreover, minimizing total error probabilities is different from minimizing total error costs. In this subjectivity, other powerful tools were adopted, such as the ROC curve representing the universe of possible events (Hanley and McNeil, 1982). In World War II, the ROC curve was first used to detect enemy objects on the battlefield. From then on, its expansion into other areas of knowledge was rapid, being widely recognized for its advantages, namely in biosciences, atmospheric forecasting or finance. The analyzes obtained through the ROC curve are considered powerful tools for validating the discriminatory power of a predictive model (Basel Committee on Banking Supervision, 2005). ROC Curve results from how the scores obtained from the prediction model are distributed between firms considered healthy and in financial distress. A perfect model would not confuse the scores between both financial health categories, but in the real world, there is an overlapping zone in which both coexist. Hence, a broad debate exists about the best cutoff point to consider in a financial distress prediction model.

This methodology, represented as a curve, is represented by an antagonistic relationship between sensitivity (the proportion of correctly classified non-failures) and specificity (the proportion of correctly classified failures) along a continuous scale of cutoff points. In other words, the area under the curve (AUC) summarizes curve performance across all thresholds, and a cutoff point is a defined criterion to separate failed from healthy firms. The greater the AUC, where x corresponds to (1-specificity) and y the sensitivity, the greater the discriminating power of the model. The ROC curve conveys the conjugation of the type I and type II error curves along an axis. In practice, AUC is a measure of prediction accuracy, where 1 will represent a perfect model. On the contrary, an AUC equal to 0.5 will demonstrate the total ineffectiveness of the model in predicting an occurrence (Altman *et al.*, 2010; Hanley and McNeil, 1982). Thus, a larger AUC indicates better predictability of the model.

The ROC curve is widespread in medical diagnosis, where there are demanding precision scales. For example, are expected AUCs between 0.80 and 0.90 for chest x-ray films and 0.80 to 0.90 for mammography. In weather forecasting, are accepted values from 0.75 for rain forecast and 0.65 for temperature intervals or fog (Swets, 1988). In one of the unavoidable references in the literature, Lemeshow *et al.* (2013) does not mention an optimal scale to describe the quality of discrimination, but in general, is used the following rule: (i) no discrimination if AUC is equal to 0.5, (ii) poor, if between 0.5 and 0.7, (iii) acceptable, if between 0.7 and 0.8, (iv) excellent, if between 0.8 and 0.9, and (v) outstanding if it is above 0.9. About financial distress prediction models in agriculture, Klepac and Hampel (2017) mentions 4 classifications: (i) eligible if AUC is between 0.50 and 0.75, (ii) good if between 0.75 and 0.92, (iii) very good if between 0.92 and 0.97, and (iv) perfect if it is above 0.97. Valaskova *et al*. (2020) defines five levels of accuracy: (i) inappropriate for bankruptcy prediction if below 0.6, (ii) poor if between 0.6 and 0.7, (iii) fair if between 0.7 and 0.8, (iv) good if between 0.8 and 0.9, and (v) excellent if above 0.9.

This study analyzes the accuracy of the presented models by examining the areas under the ROC curves and, specifically, the differences between them. Through the interaction between the different subsamples and trying out various combinations, we tested the following null hypotheses:

- Interaction between the *Global Model*[4] and the *Aggregate Models*[5]:
  H_1: Between the Global and Vineyards Models, there are no differences in the AUCs.
  H_2: Between the Global and Olive Groves Models, there are no differences in the AUCs.
  H_3: Between the Global and Portugal Models, there are no differences in the AUCs.
  H_4: Between the Global and Spain Models, there are no differences in the AUCs.
  H_5: Between the Global and Italy Models, there are no differences in the AUCs.
- Interaction between Crop Aggregates:
  H_6: Between the Vineyards and Olive Groves Models, there are no differences in the AUCs.
- Interaction between Country Aggregates:
  H_7: Between the Portugal and Spain Models, there are no differences in the AUCs.
  H_8: Between the Portugal and Italy Models, there are no differences in the AUCs.
  H_9: Between the Spain and Italy Models, there are no differences in the AUCs.
- Combined Interaction of *Individual Models*[6]:
  H_{10}: Between the Portugal Vineyards and Portugal Olive Groves Models, there are no differences in the AUCs.
  H_{11}: Between the Spain Vineyards and Spain Olive Groves Model, there are no differences in the AUCs.
  H_{12}: Between the Italy Vineyards and Italy Olive Groves Models, there are no differences in the AUCs.
  H_{13}: Between the Portugal Vineyards and Spain Vineyards Models, there are no differences in the AUCs.
  H_{14}: Between the Portugal Vineyards and Italy Vineyards Models, there are no differences in the AUCs.
  H_{15}: Between the Spain Vineyards and Italy Vineyards Models, there are no differences in the AUCs.
  H_{16}: Between the Portugal Olive Groves and Spain Olive Groves Models, there are no differences in the AUCs.
  H_{17}: Between the Portugal Olive Groves and Italy Olive Groves Models, there are no differences in the AUCs.
  H_{18}: Between the Spain Olive Groves and Italy Olive Groves Models, there are no differences in the AUCs.

3.3. *Independent variables*

Table 3 contains 12 financial ratios to be tested as potential independent variables in the model according to those most commonly present in bankruptcy and financial distress prediction studies. For this study, we combine four categories of ratios. It is in this structure that they are presented throughout this paper: (i) liquidity ratios that measure the ability of firms to honour their short-term commitments, (ii) solvency ratios /leverage that is associated with the ability to level of indebtedness and the ability to meet its payment obligations, including long-term ones, and continue to operate in the future, (iii) profitability ratios determine the ability to generate income through efficient management of resources and (iv) activity/other ratios that measure the structure of fixed assets and the operational activity of agricultural firms.

We chose to exclude financial ratios that presented inconsistent values with the expected sign in this list by logical intuition. For example, the profitability ratio that measures the relationship between earnings and equity (return on equity) could simultaneously contain negative signals in the numerator and denominator. That would result in a positive and erroneously good ratio, and we found 418 firms in this condition on our preliminary data. Also, the solvency ratio, which measures the relationship between total liabilities and equity (debt-to-equity), could be affected by negative equity found in 489 firms in our data. The result would be contrary to the perception that this ratio will worsen the greater the relationship between the numerator and denominator.

[4] The total sample, all types of crops and countries.
[5] Subsamples by type of crop, or by country.
[6] Individual interaction between crops and countries.

Table 3 - Initial set of financial ratios.

	Ratios	*Description*	*Observations*
Liquidity	CCL	Cash and equivalents / Current liabilities	Cash Ratio
	WCTA	Working capital / Total assets	
	CATA	Current assets / Total assets	
	CR	Current assets / Current liabilities	Current Ratio
Solvency/ Leverage	RETA	Retained earnings / Total assets	
	EQTA	Equity / Total assets	Shareholder Equity Ratio
	TLTA	Total liabilities / Total assets	Debt-to-Assets Ratio
Profitability	EBITTA	EBIT / Total assets	
	CFTA	(Net income + Deprec + Amortiz) / Total Assets	
	ROA	Net income / Total assets	Return on Assets
Activity/Others	STA	Sales / Total assets	Total Asset Turnover
	FATA	Fixed assets / Total assets	

Source: Own elaboration.

3.4. *Model development*

This study presents 12 models divided into two groups. The first group of six is based on aggregated data, considering all data as a whole or aggregating them according to crops or countries. The second group of six subdivides the data by countries and crops. The data considered in estimating these models are from 2018 because 2019 and 2020 only classify firms according to their financial health.

We performed a binary logistic regression, a statistical method in which several assumptions must be observed. The first is that the dependent variable is measured on a dichotomous scale. The probability of a given observation falling into one of two possible categories is predicted, healthy firm or distressed firm. The second assumption is the existence of several independent variables. The third assumption is the independence of observations, thus being mutually exclusive and exhaustive categories. Finally, the fourth assumption is that there must be a linear relationship between any continuous independent variables and the logit transformation of the dependent variable. We performed the Box-Tidwell transformation in SPSS for this last assumption, which confirmed that this assumption is not violated. The logistic model is given by:

$$Pi = \frac{1}{1 + e^{-(\beta 0 + \beta 1\ Xi1 + \beta 2\ Xi2 + \ldots + \beta nXin)}} \quad (1)$$

where, P_i = probability of financial distress, X_{ij} = j^{th} variable of the i^{th} firm, and β_j = estimated coefficient for the j^{th} variable.

3.5. *Models accuracy*

This article presents several forecasting models and analyses their explanatory power. We use the confusion matrix (Table 4) to analyse type I and II errors and the area under the ROC curve (AUC) for analyzing the occurrence of misclassifications. This matrix shows the number or percentages of false positives (FP, type I error), false negatives (FN, type II error), true positives (TP, sensitivity) and true negatives (TN, specificity). Let us assume that the *negatives* are the healthy firms and the *positives* are financially distressed firms:

The accuracy of the classification process is based on the relationship between sensitivity and specificity, according to the following equations:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

where, TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

Table 4 - Confusion matrix.

		Predicted	
		Healthy	*Financial Distressed*
Observed	*Healthy*	True negative	False positive
	Financial Distressed	False negative	True positive

Source: Own elaboration.

The AUC equation is given by:

$$AUC = \int_0^1 TPR\big(FPR^{-1}(\theta)\big)d\theta \qquad (5)$$

where, *TPR* represents the true positive rate, *FPR* = false positive rate = (1 − *Specificity*), $FPR^{-1}(\theta)$ represents the classification threshold value that corresponds to a given θ, and θ varies from 0 to 1, representing the proportion of positive samples that are correctly classified out of the total positive samples.

100(1- α)% confidence interval can be calculated using the standard normal distribution, that is:

$$AUC \pm Z_{\alpha/2} SE(AUC) \qquad (6)$$

According Hanley and McNeil (1982), the standard error of the area under the curve is given by:

$$SE(AUC) = \sqrt{\frac{AUC(1-AUC) + (n_{FD}-1)(Q1-AUC^2) + (n_H-1)(Q2-AUC^2)}{n_{FD}.n_H}} \qquad (7)$$

where, AUC = area under the ROC curve, n_{FD} = number of financial distressed firms, n_H = number of healthy firms, $Q1$ = AUC/(2-AUC), and Q2 = $2AUC^2/(1+AUC)$.

The test statistic is given as follows:

$$Z = \frac{AUC}{SE(AUC)} \qquad (8)$$

Although there is no criterion to determine the optimal cutoff for several reasons (misclassification costs, efficiency, etc.), the Youden Index (*J*) provides a criterion to determine an optimal threshold value (Fluss *et al.*, 2005), which it is maximized the equation:

$$J = Max_c(Sensitivity_c + Specificity_c - 1) \qquad (9)$$

where *c* = optimal cutoff.

In this study, for simplicity, we assume that sensitivity and specificity are equally important or desirable.

We use the same method as Hanley *et al.* (1983) to assess the differences between the AUC of the different models. This method performs a two-sided test for differences between AUCs that analyzes the proportion of positive and negative cases and the respective AUC of each model. The test returns a *p*-value determining the significance of the difference between the two curves. The statistical test is as follows:

$$z = \frac{AUC_1 - AUC_2}{\sqrt{(SE(AUC_1))^2 + (SE(AUC_2))^2 - 2.r.SE(AUC_1).SE(AUC_2)}} \qquad (10)$$

where, z = standard normal variate and r = correlation between AUCs.

4. Results and discussion

4.1. *Statistical results*

To understand how variables are revealed when forming different subsamples depending on the financial health of firms and across countries, we present the respective descriptive statistics in supplementary materials (Table S1 and Table S2). The median is the correct measure of central tendency, considering that outliers were not excluded and the sample is not uniformly distributed. As expected, and for the generality of the results, the medians are better in healthy firms, regardless of type of the crops. There are, however, some exceptions that deserve to be highlighted when the analysis considers countries. In Portugal and Italy, all ratios are consistent depending on whether firms are healthy or in financial distress. However, in Spain, CR, EQTA, and TLTA ratios present better results in Spanish financial distress firms than in healthy firms. In Portugal and Italy, all ratios are consistent depending on whether firms are healthy or in financial distress.

Confirming previous studies on the violation of the assumption of normality in the distribution of financial ratios (Deakin, 1976; Frecka and Hopwood, 1983), we performed a standard Kolmogorov-Smirnov Test, where, unsurpris-

Table 5 - Panel A: Aggregate models.

		Global Model	Crops Models		Countries Models		
			Vineyards	*Olive G.*	*Portugal*	*Spain*	*Italy*
	Constant	-1.055	-0.862	-1.085	-1.309	-1.338	-0.707
	p-Value	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
I	CCL						
	p-Value						
	WCTA						
	p-Value						
	CATA	-0.772	-0.654	-0.796			-0.985
	p-Value	(0.001)	(0.001)	(0.001)			(0.001)
	CR						
	p-Value						
II	RETA					-0.680	
	p-Value					(0.001)	
	EQTA						
	p-Value						
	TLTA	0.159			0.453		
	p-Value	(0.010)			(0.001)		
III	EBITTA	-1.352	-1.658	-1.356		-2.277	-1.587
	p-Value	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)
	CFTA						
	p-Value						
	ROA						
	p-Value						
IV	STA	-1.022	-1.646	-0.321	-2.014	-0.667	-0.980
	p-Value	(0.001)	(0.001)	(0.068)	(0.001)	(0.014)	(0.001)
	FATA						
	p-Value						
	χ^2 Model	274.614	225.754	62.495	69.785	53.647	188.385
	Model *p*-Value	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	Nagelkerke R^2	0.085	0.109	0.053	0.099	0.084	0.101
	-2 log Likelihood	4,675.28	2,881.09	1,779.06	918.90	882.22	2,772.31
	N	5,057	3,223	1,834	1,257	1,067	2,733

I (Liquidity), II (Solvency/Leverage), III (Profitability), IV (Activity/Others).
Source: Own elaboration.

ingly, we found that none of the financial ratios presents a normal distribution (Table S3 in supplementary materials).

A Spearman correlation matrix is performed to observe the correlations between covariates (Table S4 in supplementary materials). Considering that we are using a non-uniformly distributed distribution, it is preferable to the Pearson correlation matrix (Bol *et al.*, 2012). The Spearman correlation coefficient uses the order values of the observations. Thus, this coefficient is not sensitive to distribution asymmetries nor the presence of outliers, not requiring that the data come from two normal populations. Given the typology of each ratio, it is intended to select one or at most two ratios in each category. The se-

Table 6 - Panel B: Individual models.

		Vineyards			*Olive Groves*		
		Portugal	*Spain*	*Italy*	*Portugal*	*Spain*	*Italy*
	Constant	-1.695	-1.069	-0.712	-1.496	-1.638	-0.674
	p-Value	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
I	CCL						
	p-Value						
	WCTA						
	p-Value						
	CATA			-0.659			-1.159
	p-Value			(0.008)			(0.001)
	CR						
II	*p*-Value						
	RETA					-0.741	
	p-Value					(0.001)	
	EQTA						
	p-Value						
	TLTA	0.401					
	p-Value	(0.001)					
III	EBITTA			-2.579		-1.457	-0.927
	p-Value			(0.001)		(0.056)	(0.002)
	CFTA		-6.365		-2.565		
	p-Value		(0.001)		(0.023)		
	ROA						
	p-Value						
IV	STA	-1.747	-0.802	-2.240	-3.036		
	p-Value	(0.001)	(0.029)	(0.001)	(0.009)		
	FATA						
	p-Value						
	χ^2 Model	40.732	45.561	191.285	33.915	20.049	31.909
	Model *p*-Value	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	Nagelkerke R^2	0.085	0.142	0.151	0.152	0.062	0.052
	-2 log Likelihood	641.88	421.51	1,720.25	271.91	447.97	988.17
	N	855	513	1,855	402	554	878

I (Liquidity), II (Solvency/Leverage), III (Profitability), IV (Activity/Others).
Source: Own elaboration.

lection of ratios to be included in the final model goes through several combinations between variables from different categories to potentially reduce multicollinearity. In the discrimination between healthy and distressed firms, the numerical comparison is expected to be consistent with previous studies.

Performing a Mann-Whitney U-Test (Table S5 in supplementary materials), it is possible to verify that the differences in the financial ratios of healthy firms for those in financial distress are only sometimes consistent across crops or countries. Some ratios only express such differences in one of the crops (example of CCL in vineyards), and others that depend on the country (example of FATA in Olive Groves in Italy).

Table 7 - Prediction accuracy of models.

Model	*AUC*	*Cutoff*	*Confusion Matrix Parameters*				
			Accuracy	*Sensitivity*	*Specificity*	*Type I Error*	*Type II Error*
Global	0.724	0.2405	73.2%	59.3%	76.6%	40.7%	23.4%
Vineyards	0.752	0.2354	71.8%	66.3%	73.1%	33.7%	26.9%
Olive Groves	0.695	0.2397	73.8%	53.9%	78.8%	46.1%	21.2%
Portugal	0.706	0.1778	80.0%	50.6%	84.5%	49.4%	15.5%
Spain	0.694	0.2034	80.4%	49.4%	86.3%	50.6%	13.7%
Italy	0.739	0.2775	69.7%	67.5%	70.4%	32.5%	29.6%
Portugal Vineyards	0.696	0.1838	83.0%	46.2%	88.9%	53.8%	11.1%
Spain Vineyards	0.760	0.2341	80.7%	59.8%	85.0%	40.2%	15.0%
Italy Vineyards	0.788	0.2478	68.0%	79.7%	64.9%	20.3%	35.1%
Portugal Olives	0.762	0.1383	63.7%	78.4%	61.5%	21.6%	38.5%
Spain Olives	0.698	0.1627	78.2%	57.8%	81.7%	42.2%	18.3%
Italy Olives	0.669	0.3125	67.2%	58.7%	70.3%	41.3%	29.7%

AUC - Area under ROC curve.
Source: Own elaboration.

Table 8 - Comparison of differences between areas under ROC Curve.

	Difference between AUCs	*Std. error*	*z*	*p-value*
H1 Global Model ~ Vineyards	0.028	0.0156	1.793	0.0730
H2 Global Model ~ Olive Groves	0.029	0.0191	1.515	0.1298
H3 Global Model ~ Portugal	0.018	0.0255	0.705	0.4808
H4 Global Model ~ Spain	0.030	0.0258	1.164	0.2446
H5 Global Model ~ Italy	0.015	0.0157	0.958	0.3382
H6 Vineyards ~ Olives	0.058	0.0196	2.930	0.0034***
H7 Portugal ~ Spain	0.012	0.0333	0.359	0.7194
H8 Portugal ~ Italy	0.033	0.0257	1.262	0.2068
H9 Spain ~ Italy	0.045	0.0266	1.673	0.0942
H10 Portugal Vineyards ~ Portugal Olives	0.065	0.0455	1.429	0.1530
H11 Spain Vineyards ~ Spain Olives	0.062	0.0454	1.369	0.1711
H12 Italy Vineyards ~ Italy Olives	0.119	0.0246	4.838	0.0001***
H13 Portugal Vineyards ~ Spain Vineyards	0.064	0.0423	1.502	0.1330
H14 Portugal Vineyards ~ Italy Vineyards	0.092	0.0316	2.905	0.0037***
H15 Spain Vineyards ~ Italy Vineyards	0.028	0.0335	0.845	0.3979
H16 Portugal Olives ~ Spain Olives	0.064	0.0485	1.315	0.1886
H17 Portugal Olives ~ Italy Olives	0.092	0.0410	2.254	0.0242**
H18 Spain Olives ~ Italy Olives	0.029	0.0394	0.728	0.4663

*AUCs - Areas under ROC curve. ***. **. * represent .01. .05. and .10 significance levels, respectively.*
Source: Own elaboration.

For countries, and since there are three independent groups, we performed a Kruskal-Wallis Test (Table S6 in supplementary materials), which also dispenses the assumption of normality. We tested whether at least one sample comes from the same population. The null hypothesis was rejected for all financial ratios, which presupposes that there will be significant differences in the distribution of variables by country level.

The estimation of the logit model is summarised in Table 5 and Table 6. Excepted for STA in the Olive Groves aggregate model and EBITTA in the Spain Olive Groves individual model, all covariates were estimated with a p-value<0.05. However, with the p-value on the significance threshold, we chose to include them in the models as they improve the respective R^2. All estimated models present a Chi-Square goodness of fit test with an associated probability below 0.01 indicating that the current models outperform the intercept models. That is, it is concluded that the independent variables significantly influence the estimated models.

The accuracy and AUC are summarized in Table 7, which expresses the confusion matrix results.

The optimal cutoff point in these models was determined using Youden's index.

Table 8 and Figure 2 present the test results comparing the areas under the Roc curve of the different models. The main result of the differences in the areas under the curve between aggregated models is that only the Vineyards model shows differences with the Olive Groves model. Neither the global model compared with the crop or country models nor the countries themselves showed statistically significant differences. The Vineyards Model is more accurate (AUC of 0.752 against 0.695), despite the covariates chosen to be the same as the Olive Groves Model (CATA, EBITTA and STA).

Analysis of the individual models' differences results in the finding that the models sometimes present pretty significant differences. This is the case of comparing the models in Italy about Vineyards and Olive Groves. The statistical test has a p–value of less than 0.0001, the most robust rejection of the null hypothesis. In Italy, the Vineyards Model has an AUC of 0.788, which is even the best AUC of all 12 models. In turn, the Olive Groves Model from Italy has the lowest AUC of all models. In this model, the STA covariate is not included due to a lack of statistical significance, being a model with only two covariates in addition to the constant. Between different countries but with the same crops, there are also differences to be noted. The null hypothesis is also rejected in the Vineyards case between Portugal and Italy. The Italy model has the best accuracy (AUC of 0.788 against 0.696). Although both models contain the variable STA, Portugal only has two covariates, while Italy also has CATA. In the case of Olives Groves, the null hypothesis of differences between Portugal and Italy is also rejected. However, in this case, the opposite situation is registered, with Portugal registering an AUC of 0.762 while Italy is only 0.669. Interestingly, there is no covariate common to both models, highlighting that the Portuguese model uses a variable from the profitability category (CFTA) and another from the category of Activity (STA). In the case of Italy, a covariate of the liquidity category and another of profitability (EBITTA) is used.

4.2. *Discussion*

If, until now, studies dedicated to predicting bankruptcy or financial distress in agriculture generally considered agriculture as a whole, this study demonstrates that there are specificities that are not indifferent to the estimation of the models.

Although the dependent variable that determines the firm's state (healthy or distressed) is not based on variables that measure firms' activity, the STA ratio is a covariate in almost all the models presented. Only in the individual models of Spain and Italy referring to Olives Groves was this variable not shown to be statistically significant.

However, the models show a lower Nagelkerke R^2 compared to other studies. We must also consider that we did not remove outliers and limited the study to the most popular financial covariates in credit risk models. Thus, it is possible to improve the accuracy of the models by including qualitative and categorical variables. The AUCs, not stunning, can be considered eligible and suitable according to other researchers' ratings, so the models are far from useless.

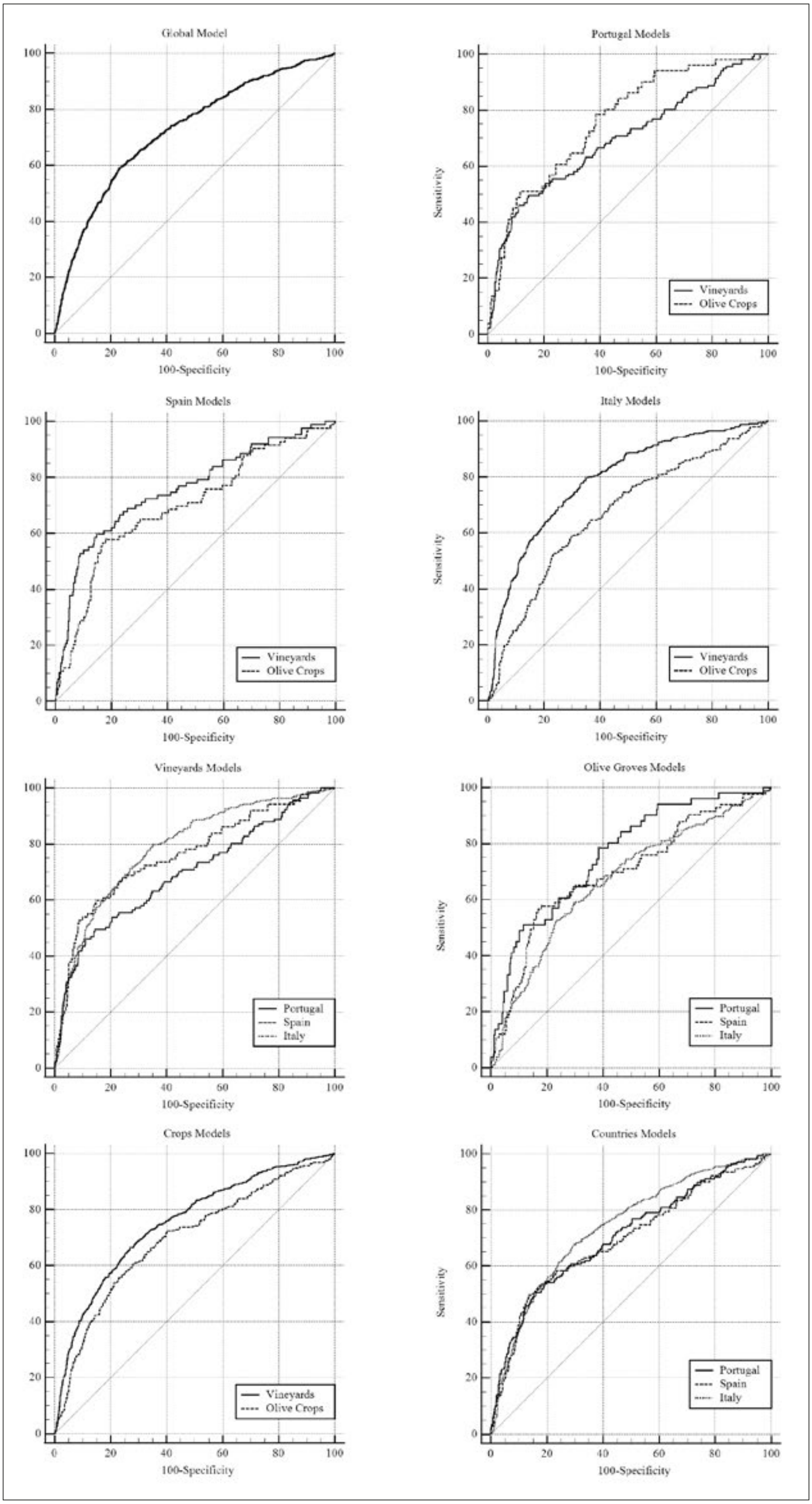


Figure 2 - Comparison of ROC curves.

Source: Own elaboration

Financial costs by an imperfect estimation of the model, namely the cost of classifying a distressed financial firm as healthy, being higher than the inverse, are the backbone of the discussion in this paper. Therefore, and since it is possible to determine a specific cutoff to separate the two categories, it is essential to have a model that, for all possible thresholds, is as accurate as possible. Therefore, the area under curves obtained from the ROC Curve is of essential importance. Moreover, in the case of agricultural activity, in which different cutoffs may be associated depending on the country or crop, it is essential to have a model that presents the best accuracy along with the possible cutoffs.

The best accuracy of the model is only sometimes consistent with the highest AUC. Aggregate models of Portugal and Spain in which the accuracy based on the confusion matrix is among the highest of the aggregate models, but which, ambiguously, have the lowest AUCs. On the contrary, the Italian model has a low accuracy compared to the other aggregate models but has the highest AUC of the country models.

In individual models, we have similar cases. The Italy Vineyards model has the lowest confusion matrix accuracy but has the highest AUC of all the individual models. On the contrary, the Portugal Vineyards model has the highest accuracy of the individual models but the lowest AUC.

In an undetermined optimal cutoff context, the AUC should be a preferable measure. However, when it is possible to determine an optimal cutoff, accuracy has the advantage of minimizing the sum of false positives and false negatives.

In comparing the accuracy of the models, we have identified statistically significant differences in the areas under the ROC curve (AUCs) for the following hypotheses, leading to the rejection of the null hypotheses:

- H_6: The comparison between Vineyards and Olives models showed a statistically significant difference in AUCs.
- H_{12}: The comparison between Italy Vineyards and Italy Olives models exhibited a highly significant difference in AUCs.
- H_{14}: The comparison between Portugal Vineyards and Italy Vineyards models demonstrated a significant difference in AUCs.
- H_{17}: The comparison between Portugal Olives and Spain Olives models resulted in a rejected hypothesis due to a significant difference in AUCs.

It is exciting that in the aggregate models, only between the aggregate model of Vineyards (0.752) against that of Olives groves (0.695), the differences are significant. All other aggregate models show no differences in accuracy. In the individual models, however, there are differences in some models, not only within the same country about crops (Italy) but also between different countries, although with the same crops. There are differences between Portugal and Italy in the AUCs, whether in the Vineyards or the Olive Groves. These results suggest creating specific models if the agriculture practised differs at the level of crops or countries.

5. Conclusion

This study is based on the estimation of forecasting models of financial difficulties in vineyards and olive groves in Portugal, Spain and Italy, Mediterranean countries with similar characteristics and agronomic practices. For this purpose, we analyzed popular financial covariates commonly used in financial distress analysis. Our variables are related to liquidity, solvency, profitability and activity of agricultural firms and are commonly used in credit risk models.

ROC curves and the corresponding areas under the curves (AUCs) allow us to conclude that, depending on the subsamples, significant differences suggest that credit risk in agriculture depends on the specifics of the agricultural activity itself. When comparing the differences in the areas under the ROC curve, we find significant variations between firms that cultivate olive groves and those that cultivate vineyards. The vineyards model is more predictive of financial distress, while the olive groves model is less accurate. However, no significant differences are observed among the various combinations of model comparisons across countries. The models that aggregate firms by country, namely Portugal, Spain, and Italy, show no significant variations. In Italy, the vineyards and olive groves models exhibit statistical differences. At the country and crop level, the difference in AUCs of firms that explore vineyards between Portugal and Italy is noticeable. This suggests that specific prediction models should be adopted in this country depending on the categorization of agricultural firms. The results also show significant differences between Spain and Italy in the case of olive groves.

This study highlights the importance of adopting region-specific predictive models for assessing credit risk in agriculture. Policymakers in Portugal, Spain, and Italy should consider the distinct characteristics of vineyards and olive groves cultivation when designing agricultural policies and financial support programs. Tailoring policies to specific crops can lead to more targeted and effective interventions to address financial distress and promote sustainable agricultural development. The distinct challenges faced by Mediterranean countries, together with the impact of climate change and agricultural trade liberalization, highlight the need for targeted interventions to address financial difficulties and improve the financial performance of rural economies in the South relative to Northern Europe. There is still to consider the likely impacts of applying the new PAC 2023-2027 and its improved sustainability measures based on environmental and climate objectives through ecological schemes. These plans, based on significant budgets, if they consider the different characteristics of agricultural firms in different countries and crops, will certainly mitigate the factors related to financial distress. Farmers and business leaders in the agricultural sector can benefit from the insights provided by the study. Understanding the differences in credit risk prediction between vineyards and olive groves cultivation can help them make more informed financial decisions and risk management strategies. Farmers need to recognise the specific factors influencing their financial health and take appropriate actions to enhance their financial sustain- ability. For the scientific community, this study highlights the importance of considering the specific characteristics of agricultural activities when developing credit risk models for the agriculture sector. This finding could prompt further research into refining and enhancing predictive models for different agricultural activities.

There are limitations to consider when interpreting the results. The first limitation is the very definition of financial distress, which determines the dependent variable of logistic regression. While bankruptcy determines the end of the firm's activity, the severity of financial distress may not put the firm in real danger. Another limitation is that there needed to be an exhaustive exploration of predictor covariates. The study focuses on the dynamics of models between different countries and crops, having selected a few potential model-independent variables. If we had access to a combination of financial variables with others that are more qualitative and even specific to agricultural activity, the predictive power of the models could be better. Also, although this study is based on a large set of data, the business structure of agriculture in these countries is complex and leaves out all farmers who are not legally constituted as a firm. This is the case of individual entrepreneurs representing a broad spectrum of family farms and other small-scale agriculture.

Future research may introduce covariates linked to the rural world, especially those specific to different crops.

References

Agarwal V., Taffler R., 2008. Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance*, 32(8): 1541-1551.

Altman E., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4): 589-609.

Altman E.I., Hotchkiss E., Wang W., 2019. *Corporate Financial Distress, Restructuring, and Bankruptcy (Analyze Leveraged Finance, Distressed Debt, and Bankruptcy)*, 4th ed. New York: John Wiley & Sons.

Altman E.I., Sabato G., Wilson N., 2010. The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk*, 6(2): 95-127.

Arnalte-Alegre E., Ortiz-Miranda D., 2013. The 'southern model' of European agriculture revisited: Continuities and dynamics. In: *Agriculture in Mediterranean Europe: Between old and new paradigms*. Bingley: Emerald Group Publishing.

Asquith P., Gertner R., Scharfstein D., 1994. Anatomy of financial distress: An examination of junk-bond issuers. *The Quarterly Journal of Economics*, 109(3): 625-658.

Basel Committee on Banking Supervision (2005). *Studies on the validation of internal rating systems*. Working Paper no. 14. Basel: Bank for International Settlements. https://www.bis.org/publ/bcbs_wp14.pdf (accessed on 12 October 2022).

Beaver W.H., 1966. Financial ratios as predictors of failure. *Journal of Accounting Research*, 4(71): 71-111.

Beopoulos N., 2017. Mediterranean agriculture in the light and shadow of the cap. In: Stathakis G., Kasimis C. (eds.), *The reform of the CAP and rural development in Southern Europe*. London: Routledge, pp. 39-51.

Bol G., Nakhaeizadeh G., Rachev S.T., Ridder T., Vollmer K.H., 2012. *Credit risk: measurement, evaluation and management*. Heidelberg: Springer Science & Business Media.

Čámská D. *et al*., 2016. Accuracy of models predicting corporate bankruptcy in a selected industry branch. *Ekonomický časopis*, 64(4): 353-366.

Caraveli H., 2000. A comparative analysis on intensification and extensification in mediterranean agriculture: dilemmas for LFAs policy. *Journal of Rural Studies*, 16(2): 231-242.

Chan M. L., Rotenberg W.D., 1988. Financial distress in the Canadian agricultural sector: A macro analysis. *Canadian Journal of Agricultural Economics/ Revue canadienne d'agroeconomie*, 36(3): 531-538.

Correia L., Behmiri N.B., Gouveia S., 2019. Drivers of wine production in the European Union: A macroeconomic perspective. *New Medit*, 18(3): 85-96.

Deakin E.B., 1976. Distributions of financial accounting ratios: some empirical evidence. *The Accounting Review*, 51(1): 90-96.

Dopuch N., Holthausen R.W., Leftwich R.W., 1987. Predicting audit qualifications with financial and market variables. *Accounting Review*, 62(3): 431-454.

FAO, 2022. *Crops and livestock products*. https://www.fao.org/faostat/en/#data/QCL (accessed on 11 February 2022).

Fitzpatrick P.J., 1934. Transitional stages of a business failure. *Accounting Review*, 9(4): 337-340.

Fluss R., Faraggi D., Reiser B., 2005. Estimation of the youden index and its associated cutoff point. *Biometrical Journal*, 47(4): 458-472.

Fraga H., Pinto J.G., Santos J.A., 2019. Climate change projections for chilling and heat forcing conditions in European vineyards and olive orchards: A multi-model assessment. *Climatic Change*, 152(1): 179-193.

Frecka T.J., Hopwood W.S., 1983. The effects of outliers on the cross-sectional distributional properties of financial ratios. *Accounting Review*, 1(58): 115-128.

Giannakis E., Bruggeman A., 2015. The highly variable economic performance of European agriculture. *Land Use Policy*, 45: 26-35.

Hanley J.A., McNeil B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1): 29-36.

Hanley J.A., McNeil B.J. *et al*., 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, 148(3): 839-843.

Hotchkiss E.S., John K., Mooradian R.M., Thorburn K.S., 2008. Bankruptcy and the resolution of financial distress. In: Eckbo B.E. (ed.), *Handbook of corporate finance*, vol. 2. Amsterdam: North-Holland, pp. 235-287.

Hsieh S.-J., 1993. A note on the optimal cutoff point in bankruptcy prediction models. *Journal of Business Finance & Accounting*, 20 (3): 457-464.

Karas M., Reznakova M., Pokorny P., 2017. Predicting bankruptcy of agriculture companies: Validating selected models. *Polish Journal of Management Studies*, 15(1): 110-120.

Khoja L., Chipulu M., Jayasekera R., 2019. Analysis of financial distress cross countries: Us- ing macroeconomic, industrial indicators and accounting data. *International Review of Financial Analysis*, 66: 101379.

Klepac V., Hampel D., 2017. Predicting financial distress of agriculture companies in EU. *Agricultural Economics – Czech*, 63(8): 347-355.

Koh H.C., 1992. The sensitivity of optimal cutoff points to misclassification costs of type i and type ii errors in the going-concern prediction context. *Journal of Business Finance & Accounting*, 19(2): 187-197.

Lemeshow S., Sturdivant R.X., Hosmer Jr D.W., 2013. *Applied logistic regression*, 3rd ed. New York: John Wiley & Sons.

Loumou A., Giourga C., 2003. Olive groves: "the life and identity of the Mediterranean". *Agriculture and human values*, 20(1): 87-95.

Mokrani A., Ben Nasr J., Sai M.B., Bachta M.S., 2022. Sustainability assessment of Tunisian olive growing systems. *New Medit*, 21(4): 3-13.

Nyitrai T., Virág M., 2019. The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences*, 67: 34-42.

Ohlson J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1): 109-131.

Platt H.D., Platt M.B., 2006. Understanding differences between financial distress and bankruptcy. *Review of Applied Economics*, 2(2): 141-157.

Platt H.D., Platt M.B., 2008. Financial distress comparison across three global regions. *Journal of Risk and Financial Management*, 1(1): 129-162.

Ruz-Carmona A., Rosa M.P., Colombo S., 2023. Quantification of production inefficiencies as a cost-savings tool for increasing the viability of traditional olive farms. *New Medit*, 22(2): 29-49.

Swets J.A., 1988. Measuring the accuracy of diagnostic systems. *Science*, 240(4857): 1285-1293.

Valaskova K., Durana P., Adamko P., Jaros J., 2020. Financial compass for Slovak enterprises: Modeling economic stability of agricultural entities. *Journal of Risk and Financial Management*, 13(5): 92.

Vavřina J., Hampel D., Janová J. *et al.*, 2013. New approaches for the financial distress classification in agribusiness. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 61(4): 1177-1182.

Whitaker R.B., 1999. The early stages of financial distress. *Journal of economics and finance*, 23(2): 123-132.

Wruck K.H., 1990. Financial distress, reorganization, and organizational efficiency. *Journal of financial economics*, 27(2): 419-444.

Appendices

Table S1 - Descriptive statistics of variables according to crops.

Type	*Ratio*	*Crops*	*Healthy Firms*						*Financial Distressed Firms*					
			N	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*	*N*	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*
I	CCL	V	2592	851.93	0.17	42728.5	-133.8	2175351	595	2.98	0.08	17.15	0.00	260.9
		OG	1438	5.33	0.25	41.2	0.00	975.4	367	6.28	0.19	48.34	0.00	749.9
	WCTA	V	2620	0.09	0.10	0.46	-9.94	1.37	603	-0.05	0.04	0.95	-18.3	1.00
		OG	1465	0.04	0.06	0.47	-5.36	1.00	369	-0.76	0.01	12.68	-243.3	0.97
	CATA	V	2620	0.45	0.40	0.29	0.00	1.00	603	0.33	0.21	0.29	0.00	1.00
		OG	1465	0.37	0.27	0.31	0.00	1.00	369	0.29	0.15	0.30	0.00	1.00
	CR	V	2592	1730.51	1.55	86133.7	-146.8	4385009	595	10.40	1.53	85.70	0.00	1983.4
		OG	1438	17.70	1.46	155.7	0.00	3616.7	367	13.73	1.08	70.58	0.00	1082.1
II	RETA	V	2620	0.18	0.19	1.27	-56.8	1.00	603	0.02	0.06	1.29	-19.57	0.99
		OG	1465	0.16	0.14	0.76	-17.6	1.00	369	-2.12	0.02	40.53	-778.3	0.99
	EQTA	V	2620	0.35	0.32	0.51	-9.94	1.00	603	0.21	0.25	1.19	-17.8	1.00
		OG	1465	0.38	0.36	0.48	-5.34	1.00	369	-0.38	0.23	12.68	-242.9	1.00
	TLTA	V	2620	0.65	0.68	0.51	0.00	10.94	603	0.79	0.75	1.19	0.00	18.79
		OG	1465	0.62	0.64	0.48	0.00	6.34	369	1.38	0.77	12.68	0.00	243.9
III	EBITTA	V	2620	0.02	0.01	0.32	-12.88	2.18	603	-0.09	-0.03	0.29	-4.27	0.61
		OG	1465	0.01	0.01	0.24	-3.27	2.30	369	-0.08	-0.03	0.31	-3.47	0.83
	CFTA	V	2620	0.04	0.03	0.32	-12.88	2.21	603	-0.07	-0.02	0.28	-4.31	0.47
		OG	1465	0.03	0.02	0.25	-4.10	2.40	369	-0.06	-0.02	0.29	-3.47	0.83
	ROA	V	2620	0.01	0.01	0.32	-12.88	2.09	603	-0.09	-0.04	0.29	-4.31	0.46
		OG	1465	0.00	0.00	0.25	-4.11	2.30	369	-0.08	-0.03	0.30	-3.47	0.83
IV	STA	V	2620	0.38	0.21	0.83	0.00	33.19	603	0.16	0.05	0.34	0.00	4.19
		OG	1465	0.30	0.11	0.53	0.00	6.74	369	0.18	0.02	0.62	0.00	7.55
	FATA	V	2620	0.55	0.60	0.29	0.00	1.00	603	0.67	0.79	0.29	0.00	1.00
		OG	1465	0.63	0.73	0.31	0.00	1.00	369	0.71	0.85	0.30	0.00	1.00

I (Liquidity), II (Solvency/Leverage), III (Profitability), IV (Activity/Others); V (Vineyards), OV (Olive Groves).
Source: Own elaboration.

Table S2 - Descriptive statistics of variables according to countries.

Type	*Ratio*	*Country*	*Healthy Firms*						*Financial Distressed Firms*					
			N	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*	*N*	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*
I	CCL	PT	1071	2057.34	0.38	66471.9	-133.8	2175351	168	4.96	0.28	23.68	0.00	260.9
		ES	892	5.20	0.39	37.96	0.00	860.5	168	9.28	0.31	59.58	0.00	749.9
		IT	2067	3.77	0.10	51.09	0.00	1699.0	626	2.69	0.07	23.33	0.00	522.2
	WCTA	PT	1089	0.16	0.20	0.53	-9.94	1.37	168	-1.49	0.10	18.80	-243.3	0.98
		ES	897	0.10	0.09	0.40	-5.36	1.00	170	-0.10	0.06	1.49	-18.32	0.96
		IT	2099	0.02	0.06	0.44	-4.68	1.00	634	-0.07	0.01	0.46	-3.19	1.00
	CATA	PT	1089	0.46	0.43	0.28	0.00	1.00	168	0.42	0.39	0.31	0.01	1.00
		ES	897	0.37	0.31	0.28	0.00	1.00	170	0.31	0.23	0.27	0.01	1.00
		IT	2099	0.42	0.33	0.32	0.00	1.00	634	0.28	0.16	0.30	0.00	1.00
	CR	PT	1071	4134.04	2.50	133991	-146.8	4385009	168	12.53	2.17	39.12	0.00	335.2
		ES	892	15.59	1.69	137.3	0.00	3616.8	168	18.48	1.77	90.47	0.01	1082.1
		IT	2067	33.61	1.26	953.5	0.00	42821.7	626	9.61	1.08	85.37	0.00	1983.4
II	RETA	PT	1089	0.23	0.27	0.69	-12.88	1.00	168	-4.90	0.02	60.05	-778.3	0.97
		ES	897	0.23	0.18	0.42	-5.39	0.99	170	-0.08	0.01	1.61	-19.57	0.99
		IT	2099	0.11	0.11	1.44	-56.77	1.00	634	0.11	0.05	0.68	-6.43	0.99
	EQTA	PT	1089	0.35	0.37	0.62	-9.94	1.00	168	-1.59	0.20	18.81	-242.9	1.00
		ES	897	0.51	0.55	0.43	-5.34	1.00	170	0.38	0.59	1.49	-17.79	1.00
		IT	2099	0.30	0.24	0.44	-8.22	1.00	634	0.30	0.19	0.40	-3.06	1.00
	TLTA	PT	1089	0.65	0.63	0.62	0.00	10.94	168	2.59	0.80	18.81	0.00	243.9
		ES	897	0.49	0.45	0.43	0.00	6.34	170	0.62	0.41	1.49	0.00	18.79
		IT	2099	0.70	0.76	0.44	0.00	9.22	634	0.70	0.81	0.40	0.00	4.06
III	EBITTA	PT	1089	0.02	0.02	0.44	-12.88	0.93	168	-0.13	-0.04	0.39	-3.47	0.53
		ES	897	0.05	0.02	0.15	-2.21	1.14	170	-0.03	-0.02	0.21	-1.73	0.65
		IT	2099	0.00	0.01	0.23	-3.41	2.30	634	-0.09	-0.03	0.28	-4.27	0.83
	CFTA	PT	1089	0.05	0.05	0.44	-12.88	0.98	168	-0.10	-0.02	0.39	-3.47	0.49
		ES	897	0.06	0.04	0.19	-4.10	1.14	170	-0.01	0.00	0.19	-1.33	0.73
		IT	2099	0.02	0.02	0.24	-3.76	2.40	634	-0.07	-0.02	0.28	-4.31	0.83
	ROA	PT	1089	0.00	0.01	0.44	-12.88	0.92	168	-0.14	-0.05	0.39	-3.47	0.46
		ES	897	0.03	0.02	0.19	-4.11	1.14	170	-0.03	-0.02	0.19	-1.38	0.63
		IT	2099	-0.01	0.00	0.24	-3.76	2.30	634	-0.09	-0.04	0.28	-4.31	0.83
IV	STA	PT	1089	0.34	0.22	0.50	0.00	6.14	168	0.17	0.07	0.25	0.00	1.63
		ES	897	0.38	0.20	0.52	0.00	6.48	170	0.27	0.08	0.54	0.00	4.19
		IT	2099	0.35	0.13	0.91	0.00	33.19	634	0.14	0.02	0.48	0.00	7.55
	FATA	PT	1089	0.54	0.57	0.28	0.00	1.00	168	0.58	0.61	0.31	0.00	0.99
		ES	897	0.63	0.69	0.28	0.00	1.00	170	0.69	0.77	0.27	0.00	0.99
		IT	2099	0.58	0.67	0.32	0.00	1.00	634	0.72	0.84	0.30	0.00	1.00

I (Liquidity), II (Solvency/Leverage), III (Profitability), IV (Activity/Others), PT (Portugal), ES (Spain), IT (Italy).
Source: Own elaboration.

Table S3 - One-sample Kolmogorov-Smirnov Test.

		Normal Parameters		*Test*	*Asymp. Sig.*
	N	*Normal Mean*	*Std. Dev.*	*Statistic*	*(2-tailed)*
CCL	4992	444.698	30789.257	0.499	0.000***
WCTA	5057	-0.001	3.468	0.386	0.000***
CA/TA	5057	0.398	0.303	0.109	0.000***
CR	4992	905.882	62066.191	0.498	0.000***
QR	4992	903.177	62066.195	0.499	0.000***
RETA	5057	-0.015	11.004	0.463	0.000***
EQTA	5057	0.288	3.479	0.419	0.000***
ICR	3710	-9157.657	312569.030	0.483	0.000***
TLTA	5057	0.712	3.479	0.419	0.000***
EBITTA	5057	-0.003	0.296	0.279	0.000***
CFTA	5057	0.016	0.300	0.290	0.000***
ROA	5057	-0.014	0.301	0.289	0.000***
ROS	4462	-13.826	523.149	0.480	0.000***
STA	5057	0.319	0.699	0.324	0.000***
FATA	5057	0.602	0.303	0.109	0.000***

Source: Own elaboration.

Table S4 - Spearman's rho coefficients.

	CCL	*WCTA*	*CATA*	*CR*	*RETA*	*EQTA*	*TLTA*	*EBITTA*	*CFTA*	*ROA*	*STA*	*FATA*
CCL	1.00											
WCTA	0.57	1.00										
CATA	0.22	0.59	1.00									
CR	0.70	0.88	0.36	1.00								
RETA	0.24	0.36	0.08	0.33	1.00							
EQTA	0.35	0.44	0.05	0.42	0.72	1.00						
TLTA	-0.35	-0.44	-0.05	-0.42	-0.72	-1.00	1.00					
EBITTA	0.18	0.28	0.25	0.18	0.40	0.27	-0.27	1.00				
CFTA	0.21	0.29	0.28	0.20	0.41	0.27	-0.27	0.91	1.00			
ROA	0.20	0.30	0.24	0.21	0.42	0.30	-0.30	0.98	0.92	1.00		
STA	0.08	0.24	0.53	0.07	0.18	0.05	-0.05	0.45	0.52	0.42	1.00	
FATA	-0.22	-0.59	-1.00	-0.36	-0.08	-0.05	0.05	-0.25	-0.28	-0.24	-0.53	1.00

For all ratios, the level of statistical significance of Spearman correlation coefficients is relevant at the 0.01 level.
Source: Own elaboration.

Table S5 - Mann-Whitney U-Test according to financial condition.

		Vineyards				Olive Groves			
		Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp.Sig. (2-tailed)*	*Mann-Whitney U*	*Wilcoxon W*	*Z*	*Asymp.Sig. (2-tailed)*
CCL	PT	36889.0	43792.0	-2.347	0.019**	8682.0	10008.0	-0.052	0.959
	ES	15421.5	19076.5	-2.070	0.038**	17343.0	127558.0	-1.583	0.113
	IT	253446.5	330867.5	-3.173	0.002***	70412.0	97673.0	-0.813	0.416
WCTA	PT	38716.0	45619.0	-1.796	0.073*	7277.0	8603.0	-2.158	0.031**
	ES	16068.0	19896.0	-1.955	0.051*	18936.0	22422.0	-0.454	0.650
	IT	249482.0	329282.0	-4.324	0.000***	66464.0	94194.0	-2.732	0.006***
CATA	PT	38170.0	45073.0	-2.016	0.044**	8503.0	9829.0	-0.577	0.564
	ES	14944.5	18772.5	-2.847	0.004***	17622.0	21108.0	-1.431	0.152
	IT	206347.0	286147.0	-8.875	0.000***	57801.0	85531.0	-5.336	0.000***
CR	PT	40615.0	47518.0	-0.828	0.408	7508.0	8834.0	-1.603	0.109
	ES	17028.0	20683.0	-0.769	0.442	18068.0	128283.0	-1.042	0.297
	IT	266428.0	343849.0	-1.778	0.075*	67224.0	94485.0	-1.798	0.072*
RETA	PT	30866.0	37769.0	-4.959	0.000***	6563.0	7889.0	-3.079	0.002***
	ES	15203.0	19031.0	-2.641	0.008***	12434.0	15920.0	-5.290	0.000***
	IT	266071.5	345871.5	-2.574	0.010**	69106.0	96836.0	-1.938	0.053*
EQTA	PT	31448.0	38351.0	-4.724	0.000***	6948.0	8274.0	-2.583	0.010**
	ES	18243.0	109194.0	-0.229	0.819	18680.0	129836.0	-0.644	0.519
	IT	278862.0	358662.0	-1.225	0.221	69389.0	97119.0	-1.853	0.064*
TLTA	PT	31448.0	304139.0	-4.724	0.000***	6948.0	68724.0	-2.583	0.010**
	ES	18250.0	22078.0	-0.223	0.824	18680.0	22166.0	-0.644	0.519
	IT	278862.0	1339558.0	-1.225	0.221	69389.0	276435.0	-1.853	0.064*
EBITTA	PT	20589.0	27492.0	-9.100	0.000***	5291.0	6617.0	-4.720	0.000***
	ES	9995.0	13823.0	-6.775	0.000***	10145.0	13631.0	-6.992	0.000***
	IT	103624.0	183424.0	-19.712	0.000***	39476.0	67206.0	-10.844	0.000***
CFTA	PT	19560.0	26463.0	-9.514	0.000***	4899.0	6225.0	-5.225	0.000***
	ES	9430.0	13258.0	-7.223	0.000***	10893.0	14379.0	-6.436	0.000***
	IT	100853.0	180653.0	-20.004	0.000***	37230.0	64960.0	-11.519	0.000***
ROA	PT	20915.0	27818.0	-8.968	0.000***	5397.0	6723.0	-4.583	0.000***
	ES	10353.0	14181.0	-6.491	0.000***	10892.0	14378.0	-6.436	0.000***
	IT	105609.0	185409.0	-19.503	0.000***	40286.0	68016.0	-10.600	0.000***
STA	PT	29582.0	36485.0	-5.477	0.000***	4979.0	6305.0	-5.146	0.000***
	ES	12173.0	16001.0	-5.046	0.000***	14291.0	17777.0	-3.909	0.000***
	IT	159662.0	239462.0	-13.810	0.000***	52596.0	80326.0	-6.947	0.000***
FATA	PT	38170.0	310861.0	-2.016	0.044**	8503.0	70279.0	-0.577	0.564
	ES	14944.5	105895.5	-2.847	0.004***	17622.0	128778.0	-1.431	0.152
	IT	206347.0	1267043.0	-8.875	0.000***	57801.0	264847.0	-5.336	0.000***

*PT (Portugal), (ES) Spain, (IT) Italy. ***, **, * represent .01, .05, and .10 significance levels, respectively.*
Source: Own elaboration.

Table S6 – Kruskal-Wallis Test[a,b].

	Global data				Vineyards				Olive Groves			
	Healthy		Fin. Distressed		Healthy		Fin. Distressed		Healthy		Fin. Distressed	
	Kruskal Wallis H	*Asymp. Sig.*	*Kruskal Wallis H*	*Asymp. Sig.*	*Kruskal Wallis H*	*Asymp. Sig.*	*Kruskal Wallis H*	*Asymp. Sig.*	*Kruskal Wallis H*	*Asymp. Sig.*	*Kruskal Wallis H*	*Asymp. Sig.*
CCL	251.706	0.000***	46.952	0.000***	163.055	0.000***	17.974	0.000***	77.390	0.000***	34.210	0.000***
WCTA	128.879	0.000***	23.143	0.000***	97.105	0.000***	15.246	0.000***	48.545	0.000***	13.816	0.001***
CATA	57.878	0.000***	38.300	0.000***	45.144	0.000***	25.661	0.000***	8.756	0.013**	14.473	0.001***
CR	157.989	0.000***	29.482	0.000***	96.376	0.000***	14.691	0.001***	76.786	0.000***	22.379	0.000***
RETA	62.581	0.000***	13.788	0.001***	37.547	0.000***	8.159	0.017**	32.836	0.000***	6.667	0.036**
EQTA	227.524	0.000***	51.655	0.000***	95.339	0.000***	23.242	0.000***	130.363	0.000***	31.017	0.000***
TLTA	227.524	0.000***	51.470	0.000***	95.339	0.000***	23.111	0.000***	130.363	0.000***	31.017	0.000***
EBITTA	98.068	0.000***	24.208	0.000***	75.460	0.000***	16.903	0.000***	41.038	0.000***	7.173	0.028**
CFTA	147.841	0.000***	37.044	0.000***	149.772	0.000***	26.618	0.000***	32.378	0.000***	12.263	0.002***
ROA	113.206	0.000***	34.662	0.000***	86.118	0.000***	23.242	0.000***	45.545	0.000***	10.149	0.006***
STA	61.860	0.000***	60.946	0.000***	75.855	0.000***	52.227	0.000***	26.313	0.000***	29.245	0.000***
FATA	57.878	0.000***	38.300	0.000***	45.144	0.000***	25.661	0.000***	8.756	0.013**	14.473	0.001***

[a.] Grouping Variable: Portugal, Spain, Italy. [b]. 2 degrees of freedom.
****. **. * represent .01. .05. and .10 significance levels, respectively.*
Source: Own elaboration.