

Universidade do Minho
Escola de Economia e Gestão

José Miguel Costa Pereira

Determinantes da Insolvência de PME
do Norte e Centro de Portugal



Universidade do Minho
Escola de Economia e Gestão

José Miguel Costa Pereira

Determinantes da Insolvência de PME do Norte e Centro de Portugal

Projeto de Mestrado
Mestrado em Finanças

Trabalho realizado sob a orientação da
Professora Doutora Sónia Silva

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositórioUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição-NãoComercial-SemDerivações
CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Dedicatória

Dedico este trabalho ao meu Pai e à minha Mãe,
por todos os sacrifícios que dia após dia fazem por mim.

Agradecimentos

Após a conclusão deste Projeto de Mestrado gostaria de deixar registados os seguintes agradecimentos. Assim, começo por agradecer aos meus pais e ao meu irmão que foram as pessoas que me permitiram chegar até aqui e que fizeram sempre de tudo para que eu conseguisse alcançar os meus objetivos. Em segundo lugar, gostaria de deixar um agradecimento muito especial à minha orientadora, Doutora Professora Sónia Maria da Silva Faria Nogueira da Silva, pela pessoa incansável que foi durante todo este processo, mostrando-se sempre disponível para qualquer esclarecimento e ajuda e pelo apoio demonstrado a todos os níveis ao longo desta caminhada. Gostaria também de agradecer aos meus Professores e colegas de Mestrado envolvidos neste processo, que foi sem dúvida um dos maiores desafios da minha vida até à data. Agradecer, também, a todos os meus amigos e colegas de trabalho que se têm cruzado comigo, ao longo destes anos, e que permitiram que de uma forma ou de outra o meu objetivo fosse concretizado. Por último, mas não menos importante deixar uma palavra de apreço à nBanks pela oportunidade que me deu de realizar este projeto.

Desta forma, deixo aqui o meu sincero obrigado, pelo carinho, compreensão e força que me deram, a todos aqueles, em geral, que fizeram parte deste trajeto.

Declaração de Integridade

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducentes à sua elaboração.

Mais declaro que conheço e respeitei o Código de Conduta Ética da Universidade do Minho.

Resumo

As pequenas e médias empresas representam mais de 99% do tecido empresarial Português, sendo que a sua principal fonte de financiamento é o crédito bancário.

Neste contexto torna-se importante que os agentes económicos disponham de instrumentos de apoio à gestão e de avaliação da situação financeira das empresas, de forma a anteciparem cenários e traçarem uma estratégia para a empresa.

Assim, o objetivo deste projeto é desenvolver um modelo empírico que identifique quais os fatores considerados determinantes para o processo de insolvência. Para o efeito, foi analisado um painel longitudinal de dados recorrendo a diferentes metodologias, entre as quais, a Análise Discriminante Multivariada, a Análise Logística e a Análise de *Matching*.

A amostra é composta por um grupo de empresas insolventes e por um grupo de empresas saudáveis do Norte e Centro de Portugal, transversal a todos os setores de atividade, exceto setor financeiro e setores altamente regulados como a energia. Os dados utilizados estão compreendidos entre 2011 e 2019 e foram recolhidos da base de dados ORBIS da Bureau van Dijk.

As variáveis selecionadas para o modelo final são as seguintes: Estrutura de Endividamento, Margem EBIT, Rácio Passivo – Capital Próprio, Rácio de Cobertura de Juros, Rotação do Ativo, Dimensão e Idade. O modelo final foi sujeito a vários testes de robustez, sendo que a capacidade preditiva deste modelo prevê corretamente mais de 64% das observações. O modelo final assume uma maior capacidade preditiva quando testado para dois e três anos antes da insolvência, prevendo corretamente entre 80 a 90% das observações.

Palavras-chave: Análise Discriminante Multivariada, Insolvência, Modelos de Previsão de Insolvência, Pequenas e Médias Empresas, *Propensity Score Matching*.

Abstract

Small and medium-sized companies represent more than 99% of the Portuguese business ecosystem whereas their main source of financing is bank credit.

In this context, it is important that economic agents have at their disposal instruments to support and assess the financial situation of companies, to anticipate scenarios and be able to draw a strategy for their organizations.

Thus, this project aims to develop an empirical model that identifies which factors are considered as determinants for the insolvency process. For this purpose, a longitudinal panel of data was analysed using different methodologies, including Multivariate Discriminant Analysis, Logistic Analysis and Matching Analysis.

The sample is composed of a group of insolvent companies and a group of healthy companies from the North and Center of Portugal from all sectors of activity, except for the financial sector and highly regulated sectors such as energy. The data covers the 2011-2019 period and was collected from the Bureau van Dijk's ORBIS database.

The variables selected for the final model are the following: Debt Structure, EBIT Margin, Debt-to-Equity Ratio, Interest Coverage Ratio, Asset Turnover, Size and Age. The final model was subjected to several robustness tests, presenting an overall forecasting accuracy of more than 64% of the observations. The final model assumes a higher forecasting capacity when tested for two and three years before the insolvency event, correctly predicting between 80 and 90% of observations.

Keywords: Insolvency, Insolvency Prediction Models, Multivariate Discriminant Analysis, Propensity Score Matching, Small and Medium Enterprises.

Índice

Resumo	vi
Abstract	vii
Índice de Tabelas	x
Índice de Gráficos	xi
Lista de Abreviaturas	xii
1. Introdução	1
2. Enquadramento teórico	5
2.1. Insolvência e PME	5
2.1.1. Conceito de Insolvência	5
2.1.2. Conceito de PME	6
2.2. Modelos de Previsão de Insolvência	6
2.2.1. Modelos gerais	7
2.2.2. Modelos específicos para PME	13
2.3. Objetivos do estudo	16
3. Apresentação da Entidade Parceira	17
4. Metodologia	19
4.1. Descrição da Metodologia	19
5. Descrição da Amostra e das Variáveis	22
5.1. Amostra	22
5.2. Variáveis	23
5.2.1. Variável Dependente	23
5.2.2. Variáveis Explicativas	23
5.3. Estatísticas Descritivas	27
5.4. Análise Univariada	32
6. Análise Multivariada	35

6.1. Análise Discriminante Múltipla	35
6.2. Análise Logística	42
6.3. Propensity Score Matching	48
6.4. Modelo Final.....	53
6.4.1. Resultados do Modelo <i>Logit</i> para dois e três anos antes da Insolvência	53
6.4.2. Resultados do Modelo <i>Probit</i> para dois e três anos antes da Insolvência.....	56
6.4.3. Modelo Final de Previsão de Insolvência	57
6.4.4. Teste de Robustez	59
6.4.5. Aplicação prática do modelo	61
7. Conclusão.....	63
Referências bibliográficas	65
Apêndice.....	67

Índice de Tabelas

Tabela 1 – Definição de PME	6
Tabela 2 – Resultados de Ohlson (1980)	12
Tabela 3 – Modelo de Altman e Sabato (2007)	15
Tabela 4 – Variáveis contruídas para o estudo.....	24
Tabela 5 – Estatísticas Descritivas: Amostra Total	28
Tabela 6 – Estatísticas Descritivas por NUTS2	29
Tabela 7 – Estatísticas Descritivas por <i>Status</i>	30
Tabela 8 – Descrição da amostra por ano e <i>Status</i>	31
Tabela 9 – Descrição da amostra por tipo de indústria e por região	31
Tabela 10 – Descrição da amostra por região e por <i>Status</i>	32
Tabela 11 – Teste às médias e às medianas em função das NUTS2 e do <i>Status</i>	33
Tabela 12 – Classificação das empresas por <i>Status</i>	36
Tabela 13 – Erros tipos I e II.....	36
Tabela 14 – Coeficientes da Função Discriminante	37
Tabela 15 – Função Discriminante por grupo.....	39
Tabela 16 – <i>Structure Matrix</i>	40
Tabela 17 – Análise ANOVA.....	41
Tabela 18 – Resultados antes do Processo de Estandarização das Variáveis.....	44
Tabela 19 – Resultados após o Processo de Estandarização das Variáveis.....	46
Tabela 20 – Estimação do <i>Propensity Score</i>	49
Tabela 21 – Resultados do teste de qualidade do <i>matching</i>	50
Tabela 22 – Estimação do <i>Propensity Score</i> para um modelo restrito	51
Tabela 23 – Resultados do teste de qualidade do <i>matching</i>	52
Tabela 24 – <i>Goodness-of-fit</i> do <i>propensity score</i>	53
Tabela 25 – Resultados da estimação do modelo <i>Logit</i>	54
Tabela 26 – Resultados da estimação do modelo <i>Logit</i> incluindo controlo para efeitos temporais, localização e indústria.....	55
Tabela 27 – Resultados da estimação do modelo <i>Probit</i>	57
Tabela 28 – Resultados da estimação do modelo <i>Logit</i> final	58
Tabela 29 – Resultados da estimação <i>Out-of-Sample</i>	60
Tabela A – Definição das indústrias utilizadas no estudo.....	67

Índice de Gráficos

Gráfico 1 – Coeficientes <i>standardized</i>	38
Gráfico 2 – ROC (<i>Receiver Operating Characteristic Curve</i>)	43
Gráfico 3 – ROC (<i>Receiver Operating Characteristic Curve</i>)	45
Gráfico 4 – Significância dos Coeficientes Estandarizados	47
Gráfico 5 – Curva ROC (Modelo Final)	59
Gráfico 6 – Curva ROC (<i>Análise Out-of-Sample</i>)	61

Lista de Abreviaturas

ADM – Análise Discriminante Múltipla

CAE – Classificação das Atividade Económicas

CIRE – Código da Insolvência e da Recuperação de Empresas

NUTS – Nomenclatura das Unidades Territoriais para Fins Estatísticos

PME – Pequenas e Médias Empresas

PSM – *Propensity Score Matching*

1. Introdução

Segundo dados recolhidos da Pordata¹, o tecido empresarial português era composto, em 2018, maioritariamente por PME (pequenas e médias empresas), mais precisamente, 99,9% das empresas portuguesas eram consideradas micro, pequenas ou médias. Em relação, à estrutura empresarial Portuguesa pode ainda dizer-se que, apesar das PME fazerem parte da maioria, apenas representam, aproximadamente, 60% do volume de negócios do setor empresarial português.

Tendo em conta o peso que assumem na economia portuguesa, torna-se cada vez mais pertinente analisar eventos recorrentes aos quais estas empresas estão mais propensas, como é o caso da situação de insolvência. A análise da insolvência é uma questão que tem vindo a ganhar preponderância em Portugal, ao longo dos últimos anos, pelo seu impacto socioeconómico. O número de insolvências está altamente correlacionado com o ciclo económico, sendo que nos últimos anos o número de abertura de processos judiciais de insolvência tem vindo a diminuir acompanhando pela recuperação económica sentida desde a alteração da política monetária conduzida pelo Banco Central Europeu na segunda metade de 2013. Além da relação entre a insolvência das PME e os ciclos económicos, existe também uma relação entre a insolvência das PME e fatores como a má gestão financeira e organizacional, a dimensão e a idade da organização (Sousa & Oliveira, 2014).

A crise económica e financeira de 2008, conhecida como crise do crédito *subprime* devido ao facto de o financiamento concedido estar exposto a um elevado risco de incumprimento, tornou ainda mais restrito o acesso ao financiamento externo por parte das empresas, o que, por conseguinte, conduziu muitas empresas portuguesas a uma situação de insolvência e que muitas outras passassem por dificuldades financeiras que, ainda hoje, se fazem sentir, agravando ainda mais os seus problemas de competitividade perante os seus concorrentes internacionais. Este evento tornou-se, particularmente, relevante para a subsistência das PME, uma vez que estas dependem, em grande parte, deste tipo de financiamento para fazer face a necessidades de financiamento de curto e longo prazo. No entanto, atualmente, as PME portuguesas têm-se preparado, cada vez mais, para este tipo de fenómenos, apostando no reforço da competitividade, através de ligações frequentes com instituições de ensino superior e apostando, fortemente, no marketing e na transformação digital.

¹ A Pordata é uma base de dados sobre Portugal com estatísticas oficiais e certificadas sobre o país e a Europa, dividida num amplo conjunto de temas como a população, educação, saúde, entre outros. Disponível em: www.pordata.pt, acedido a 26/09/2020.

O estudo de métodos ou modelos de previsão de falência ou insolvência tem sido feito ao longo de décadas e muitos têm sido os autores a investigar sobre esta temática, criando modelos de previsão com um elevado poder explicativo da situação de insolvência, através da análise de rácios económicos e financeiros. Os estudos iniciais captavam, através da análise de um conjunto de rácios financeiros, os determinantes de insolvência examinados de forma isolada. Posteriormente, estudos mais recentes, apontam que apenas a utilização de indicadores financeiros isolados não é suficiente para apurar o impacto que outro tipo de informações, como, por exemplo, a idade, a dimensão ou a localização da empresa possam ter na previsão de insolvência. Por outro lado, variáveis de natureza macroeconómica também têm um grande impacto, uma vez que, ciclos económicos recessivos conduzem a um aumento de empresas em dificuldades financeiras, ao contrário do que acontece em períodos de crescimento económico.

Os autores pioneiros neste tipo de investigação foram Beaver (1966) e Altman (1968), que utilizaram rácios/indicadores financeiros e económicos para desenvolver modelos univariados e multivariados de previsão de falência.

Na sua investigação, Beaver (1966) utilizou diversos indicadores financeiros através de duas abordagens. A primeira baseou-se numa classificação entre empresas com elevada e baixa probabilidade de falência e, para tal, analisou o comportamento dos indicadores financeiros até cinco anos anteriores às falências. A segunda baseou-se na metodologia *Bayesiana* que consiste em observar os indicadores em períodos anteriores à falência. O autor concluiu que o poder preditivo de cada variável é bastante distinto, destacando-se a relação entre fluxo de caixa operacional bruto e dívida bruta.

Altman (1968) adotou a análise discriminante multivariada para selecionar os rácios com melhor capacidade preditiva da situação de insolvência alguns anos antes dessa situação. O autor propôs um modelo com 5 rácios - o *Z-score* - que apesar de ter sido desenvolvido há várias décadas e de existirem vários outros modelos de previsão de insolvência, continua a ser a principal ferramenta de previsão de falência e de dificuldades financeiras em todo o mundo (Altman, Iwanicz-Drozdowska, Laitinen & Suvas, 2017).

Posteriormente, ao trabalho pioneiro de Beaver (1966) e Altman (1968), vários métodos têm sido utilizados em diferentes investigações que têm sido estudadas, principalmente, nos países industrializados, como, por exemplo, em Inglaterra (Taffler & Tisshaw, 1977), em França (Altman, Margaine, Schlosser & Vernimmen, 1974), entre outros países da Europa, como a Bélgica e os Países Baixos (Abrahamse & Frederikslust, 1975).

Mais tarde, Ohlson (1980) apresenta um modelo de regressão logística, que permite um aumento da eficiência do modelo. Para o cálculo da probabilidade de insolvência, entre 0 e 100%, o autor utiliza a regressão logística, tornando este modelo mais simples e fácil de interpretar em relação aos anteriores.

No caso das PME, mais especificamente, estas diferem das grandes empresas, uma vez que necessitam de modelos individuais e específicos de previsão de falência. São vários os autores que estudam este caso concreto recorrendo à análise discriminante multivariada, entre os quais: Edmister (1972); Altman e Sabato (2007) e Altman, Sabato e Wilson (2010). O primeiro estudo focado na previsão de falências deste tipo de empresas foi realizado por Edmister (1972), através da análise de rácios financeiros. O autor realça a importância da seleção dos rácios para se conseguir chegar a um modelo eficaz de previsão de falência.

Neste contexto, a presente investigação tem como principal objetivo identificar os fatores considerados determinantes na previsão da situação de insolvência de PME portuguesas, nomeadamente do Norte e Centro de Portugal. Este trabalho resulta de uma parceria entre a Escola de Economia e Gestão da Universidade do Minho e a *fintech* nBanks, que tem como principal cliente as PME portuguesas. Por conseguinte, os resultados deste estudo têm como propósito serem aplicados aos serviços oferecidos pela nBanks. Assim sendo, a capacidade de previsão do modelo que reúne as determinantes de insolvência assume grande relevância, uma vez que através de uma análise antecipada do desempenho das empresas, tanto a nível financeiro como económico, permite à nBanks estabelecer padrões, antecipar cenários e prevenir situações de insolvência, otimizando assim os seus serviços.

Este estudo tem por base uma amostra que cobre o período temporal entre 2011 a 2019. A informação económica e financeira das empresas incluídas neste projeto foi recolhida na base de dados Orbis da Bureau van Dijk para todos os setores de atividade, com exceção do setor público, do setor financeiro e de atividades altamente reguladas como a água e a energia, para empresas que entraram em insolvência e para empresas que permaneceram ativas, do Norte e Centro de Portugal, de acordo com a partição feita a partir da classificação territorial NUTS2.

Seguindo os estudos empíricos de Altman (1968) e de Ohlson (1980), os dados foram analisados recorrendo às metodologias da Análise Discriminante Múltipla (ADM) tal como em Altman (1968), de forma a se selecionarem as variáveis que melhor se ajustam à construção de um modelo logístico de previsão de insolvência à semelhança do modelo de Ohlson (1980). Será também aplicada a técnica de *Propensity Score Matching* (PSM), proposta por Rosenbaum and

Rubin (1983), através da seleção de dois grupos de empresas comparáveis (um grupo de empresas de tratamento - declaradas insolventes - e um grupo de controlo - de empresas declaradas ativas), antes da situação de insolvência, em termos de liquidez, rentabilidade, endividamento, setor de atividade, dimensão ou idade.

O modelo final de previsão de insolvência proposto neste estudo inclui as seguintes variáveis: Estrutura de Endividamento, Margem EBIT, Rácio Passivo – Capital Próprio, Rácio de Cobertura de Juro, Rotação do Ativo, Dimensão e Idade, e apresenta uma capacidade preditiva de mais de 64% de observações corretamente classificadas.

Este trabalho divide-se nos seguintes sete capítulos: no Capítulo 2 é feita a contextualização do tema do projeto, onde é apresentada a base teórica que suporta o presente estudo; no Capítulo 3 é apresentada a entidade parceira do estudo em causa; no Capítulo 4 são descritas as metodologias adotadas para o estudo; no Capítulo 5 são descritas a amostra e as variáveis; no Capítulo 6 é apresentada a estimação do modelo de previsão de insolvência e é feita a análise dos resultados obtidos; no Capítulo 7 são apresentadas as conclusões resultantes do presente estudo.

2. Enquadramento teórico

2.1. Insolvência e PME

2.1.1. Conceito de Insolvência

Em primeiro lugar, é conveniente fazer a distinção entre os termos “falência” e “insolvência”, uma vez que estes são confundidos frequente e facilmente, pelo senso comum. No entanto, existem algumas diferenças que os distinguem, tornando-se importante perceber os dois conceitos. A insolvência foca-se na análise dos conceitos legais, definindo-se como insolventes as empresas que deram início aos trâmites jurídicos com vista à sua liquidação, por sua vez, a falência é associada à incapacidade que ocorre quando uma empresa não é capaz de fazer face aos seus compromissos financeiros com os bens que tem ao dispor (Sousa & Oliveira, 2014).

A literatura não é consensual em relação ao conceito de insolvência, utilizando diferentes termos para se referir a este. Segundo Beaver (1966, p.71), “falência é definida como a incapacidade de uma empresa de pagar as suas obrigações financeiras. Operacionalmente, diz-se que uma empresa falhou quando qualquer um dos seguintes eventos ocorre: *bankruptcy, bond default, an overdrawn bank account, or nonpayment of a preferred stock dividend.*”. Mais tarde, Ohlson (1980, p.111) considera que “não há consenso sobre o que constitui “*failure*”, sendo que as definições variam significativamente e arbitrariamente entre estudos”. Altman (1993) considera que a insolvência se deve a uma performance negativa, que acontece quando uma empresa não é capaz de cumprir os seus compromissos, evidenciando problemas financeiros que poderão ser temporários, não sendo garantido que esta encerre portas. Já Altman e Hotchkiss (1993) utilizam diferentes termos, para se referirem ao insucesso das empresas, como por exemplo: i) *failure*, que não implica a extinção da empresa, uma vez que, estando a mesma em falência económica, pode durante um certo período conseguir cumprir as suas obrigações, assim quando a empresa deixa de conseguir fazer face às suas obrigações entra em *legal failure*; ii) *insolvency*, considerando que a empresa é insolvente quando o valor do passivo total é excessivamente superior ao valor do ativo total; iii) *default*, acontece quando uma empresa não cumpre as suas obrigações contratuais perante a entidade credora, podendo gerar uma ação legal contra a primeira; por último, iv) *bankruptcy*, consiste numa situação de insolvência permanente, sendo o mais comum em relação a este termo, a insolvência ser declarada formalmente por um tribunal. Em, Portugal, e de acordo com o Código da Insolvência e da Recuperação de Empresas (CIRE), a insolvência é decretada quando as responsabilidades vencidas são superiores à massa insolvente. Esta será a definição adotada doravante neste trabalho.

2.1.2. Conceito de PME

No que diz respeito ao conceito de PME não existe uma definição comum quando a análise é feita para diferentes países. Os critérios que distinguem as pequenas e médias empresas variam de país para país, tendo em conta fatores quantitativos como, por exemplo, o volume de negócios, o total do ativo e o número de empregados, e fatores qualitativos como a organização do trabalho e o tipo de indústria. Para o presente estudo, a definição de PME a ser utilizada terá como base os critérios previstos – Recomendação 2003/361/CE de 6 de maio - pela União Europeia.

Assim, segundo o Decreto-Lei n.º 372/2007, de 6 de novembro, para uma empresa ser classificada como PME deverá cumprir 3 requisitos, entre os quais: empregar menos de 250 trabalhadores, o volume de negócios não exceder 50 milhões de euros ou o balanço total não exceder 43 milhões de euros. Em Portugal, as PME assumem particular relevância, pois representam 99,9% do tecido empresarial português. A Tabela 1 representa os fatores de classificação das PME de acordo com cada categoria de dimensão.

Tabela 1 – Definição de PME

Dimensão da empresa	Nº de Empregados	Volume de Negócios	Balanço Total
Média	<250	≤50 milhões de EUR	≤43 milhões de EUR
Pequena	<50	≤10 milhões de EUR	≤10 milhões de EUR
Micro	<10	≤2 milhões de EUR	≤2 milhões de EUR

Fonte: Adaptado do IAPMEI: <https://www.iapmei.pt/PRODUTOS-E-SERVICOS/Qualificacao-Certificacao/Certificacao-PME.aspx>

2.2. Modelos de Previsão de Insolvência

O estudo da previsão de insolvência tem vindo a ser feito ao longo de décadas e muitos têm sido os autores a investigar sobre esta temática através de vários métodos, o que contribuiu para a extensa literatura relacionada com os modelos de previsão de insolvência. O interesse pelo estudo de modelos de previsão de insolvência deve-se ao facto da possível aplicação destes a contextos reais, pelos diversos agentes económicos. Estes modelos, para além de preverem, em termos financeiros, o desempenho futuro de uma determinada organização, auxiliam na tomada de decisão e permitem perceber o que não está a funcionar bem dentro de uma organização, e assim reagir mais rapidamente a possíveis problemas.

Desta forma, serão apresentados nos Subcapítulos seguintes os estudos mais relevantes e que mais contribuíram para a presente investigação, que têm em conta modelos gerais de

previsão de insolvência e modelos específicos para pequenas e médias empresas.

2.2.1. Modelos gerais

Os autores pioneiros no estudo de previsão de insolvência foram Beaver (1966) e Altman (1968) que desenvolveram estudos baseados na análise univariada e multivariada, respetivamente, para prever o estado de falência de uma dada empresa através de um grupo de rácios financeiros. A análise univariada serviu de suporte às primeiras investigações de autores como FitzPatrick (1932), Smith e Winakor (1935), Merwin (1942) e Tamari (1966), que analisaram os rácios de previsão de insolvência. Desta forma, Beaver (1966) é conhecido como um dos pioneiros no estudo da previsão de insolvência.

Em 1966, Beaver desenvolveu um estudo de previsão com recurso à análise univariada, que consiste na análise isolada dos vários rácios e na comparação de rácios entre empresas insolventes e empresas ativas. O autor analisou uma amostra composta por empresas insolventes e pelas suas *peers* saudáveis, com base na sua dimensão, com o objetivo de determinar quais os rácios com maior capacidade preditiva. O estudo continha uma amostra de 158 empresas dos Estados Unidos das quais 79 insolventes e 79 saudáveis, para um período de 10 anos (1954-1964) e foram utilizados cerca de 30 rácios divididos em 6 categorias. De referir que os dados relativos às empresas insolventes remetem até 5 anos antes do ano de declaração de insolvência. O autor selecionou as empresas a estudar através de um processo de *matching* por indústria, dimensão e região, selecionando apenas o rácio de cada categoria com maior capacidade preditiva. Foram testados os seguintes 6 rácios financeiros:

- *Cash Flow/Total Debt;*
- *Net Income/Total Assets;*
- *Total Debt/Total Assets;*
- *Working Capital/Total Assets;*
- *Current Assets/Current Liabilities;*
- *No-credit interval.*

Os rácios foram utilizados para ordenar as empresas por ordem crescente de modo a definir o ponto de *cutoff* para cada um destes. O autor concluiu que o rácio com maior poder

explicativo foi o rácio *Cash Flow/Total Debt*, apresentando valores entre 87% e 78% para 1 e 5 anos antes da insolvência, respetivamente. O estudo demonstrou que a análise de rácios é útil no que concerne à previsão de insolvência, no entanto, é necessário ter em consideração que nem todos os rácios apresentam um poder preditivo semelhante. Os resultados sugeriram ainda, que à medida que se aproximavam do ano da falência, a média dos rácios das empresas que entraram em falência piorava, o que não acontecia com as empresas saudáveis. No entanto, este tipo de análise levantou críticas devido ao facto de a análise individual e isolada de um rácio não conter informação suficiente para prever a falência da empresa como um todo.

Com o objetivo de contornar algumas lacunas associadas à análise univariada, Altman (1968) recorreu à análise discriminante multivariada (ADM), para desenvolver um modelo de previsão de insolvência denominado por *Z-score*. O autor considerava que os estudos realizados até à data com recurso a análises univariadas, apesar de serem relevantes do ponto de vista da análise individual de rácios, apresentavam limitações do ponto de vista global da situação financeira de uma empresa. A sua amostra continha 66 empresas dos Estados Unidos (33 insolventes e 33 saudáveis), com um ativo total superior a 1 milhão de dólares, para o período temporal de 1946 a 1965. Foram examinados 22 potenciais rácios e selecionados 5 que, quando combinados, demonstraram ter a maior capacidade de previsão de falência. As variáveis foram divididas nas seguintes cinco categorias: liquidez, rentabilidade, alavancagem, cobertura e atividade. A função que resultou da análise discriminante foi a seguinte:

$$Z = 0,012X_1 + 0,014X_2 + 0,033X_3 + 0,006X_4 + 0,999X_5$$

Onde:

$$X_1 = \textit{Working Capital/Total Assets}$$

$$X_2 = \textit{Retained Earnings/Total Assets}$$

$$X_3 = \textit{EBIT/Total Assets}$$

$$X_4 = \textit{Market Value Equity/Book Value Total Debts}$$

$$X_5 = \textit{Sales/Total Assets}$$

X_1 : O fundo de maneio permite perceber o nível de liquidez da empresa. De todos os rácios de liquidez este é o mais significativo para o modelo, de acordo com o autor.

X_2 : Este rácio reflete os lucros que transitaram do ano n para o ano $n+1$ para desenvolver a atividade da empresa, estando implícita a idade da empresa. Para o autor, uma empresa mais jovem provavelmente apresenta menores resultados transitados do que uma empresa mais antiga. Desta forma, algumas empresas poderão ser classificadas de forma errada por serem mais recentes.

X_3 : De acordo com o autor, este rácio é o que melhor estima a rendibilidade dos ativos da empresa, não incluindo os efeitos fiscais. Uma vez que a existência de uma empresa depende, em grande parte, da capacidade de rentabilizar os seus ativos, Altman (1968) considera este rácio fundamental para os estudos de previsão de insolvência, uma vez que este rácio é o que mais contribui para a categorização entre grupos (insolvente *versus* saudável).

X_4 : Este rácio tem a capacidade de medir a potencial desvalorização dos ativos face aos passivos, levando assim a empresa à insolvência.

X_5 : Podemos definir este rácio como a rotação do ativo, que demonstra a capacidade de uma empresa gerar vendas com base nos seus ativos. Apesar deste rácio apresentar um menor grau de significância, a sua importância deve-se à sua relação com as outras variáveis do modelo.

A função *Z-Score* pode ser interpretada da seguinte forma:

- $Z > 2,99$ – A probabilidade de falência é reduzida;
- $1,8 < Z < 2,99$ – Incerteza quanto à possibilidade de a empresa entrar em falência;
- $Z < 1,80$ – A probabilidade de falência é elevada.

Este estudo demonstrou que a análise discriminante multivariada é estatisticamente mais significativa do que a análise univariada. O modelo *Z-score* de Altman (1968) apresenta uma capacidade de previsão de falência de 95% um ano antes desta situação acontecer. Em relação à capacidade preditiva para 2, 3, 4 e 5 anos antes da falência, o modelo apresentou as seguintes capacidades de previsão, 72%, 48%, 29% e 36%, respetivamente.

Mais tarde Altman, Haldeman e Narayanan (1977), através de algumas adaptações ao modelo *Z-score*, desenvolveram um novo modelo denominado ZETA, mais adequado ao contexto

empresarial da época. Para este modelo, os autores utilizaram uma amostra composta por 111 empresas dos Estados Unidos do setor industrial e do retalho, das quais 53 eram insolventes e 58 saudáveis, para o período entre 1965 a 1972, sendo que neste estudo foi analisado o impacto da dimensão nas alterações estruturais da empresa.

A função discriminante ZETA era constituída pelos seguintes 7 rácios:

$$X_1 = EBIT/Total Assets$$

$$X_2 = Standard\ error\ of\ estimate\ around\ a\ ten\ year\ trend\ in\ X_1$$

$$X_3 = \log (EBIT/Total\ Interest\ Payments)$$

$$X_4 = Retained\ Earnings/Total\ Assets$$

$$X_5 = Current\ Assets/Current\ Liabilities$$

$$X_6 = Common\ Equity/Total\ Capital$$

$$X_7 = \log (Total\ Assets)$$

Os autores concluíram que a capacidade preditiva deste modelo era mais elevada na classificação das empresas insolventes dois a cinco anos antes da insolvência. Em relação ao ano anterior à insolvência, os resultados são semelhantes aos do modelo *Z-score*. Apresentou uma capacidade de precisão de 90% para um ano antes e cerca de 70% até cinco anos antes da insolvência. Por conseguinte, concluiu-se novamente que a análise discriminante multivariada apresenta resultados estatisticamente mais significativos do que uma abordagem univariada.

Verificou-se também que uma das limitações do modelo *Z-score* de Altman (1968) era o facto de este não permitir ser aplicado em empresas não cotadas, dada a sua incapacidade e dificuldade em estimar o valor de mercado dos capitais próprios. Assim, Altman (1983) alterou o rácio X_4 , substituindo o valor de mercado dos capitais próprios pelo seu valor contabilístico.

Deste modo, a função discriminante para empresas não cotadas é a seguinte:

$$Z = 0,717X_1 + 0,847X_2 + 3,107X_3 + 0,420X_4 + 0,998X_5$$

Onde se destaca a diferença em X_4 :

$$X_1 = Working\ Capital/Total\ Assets$$

$$X_2 = \text{Retained Earnings/Total Assets}$$

$$X_3 = \text{EBIT/Total Assets}$$

$$X_4 = \text{Book Value Equity/Total Liabilities}$$

$$X_5 = \text{Sales/Total Assets}$$

Os novos resultados geraram uma ligeira alteração nos grupos de classificação relativos aos valores do índice *Z-score*, sendo a interpretação do coeficiente efetuada tendo em conta os seguintes intervalos:

- $Z > 2,90$ – A probabilidade de falência é reduzida;
- $1,23 < Z < 2,90$ – Incerteza quanto à possibilidade de a empresa entrar em falência;
- $Z < 1,23$ – Zona de Perigo – A probabilidade de falência é elevada.

Apesar do estudo de Altman (1968) apresentar resultados positivos, alguns autores (p.e., Ohlson, 1980) criticam a escolha da Análise Discriminante Multivariada (ADM) devido à constante violação de alguns pressupostos implícitos neste tipo de análise, nomeadamente, no que respeita às variáveis independentes incluídas no modelo que devem apresentar uma distribuição normal e que as matrizes de variância e covariância dos dois grupos de empresas (falidas e não falidas), deverão ser iguais.

Por conseguinte, Ohlson (1980) considerando as limitações associadas ao uso da ADM, aplicou, pela primeira vez, como alternativa ao tipo de análise anterior, a regressão logística ao estudo da previsão de insolvência. Na prática, a Análise *Logit* não requer os pressupostos restritivos exigidos na ADM, o que permite trabalhar com amostras desproporcionais e aumentar a validade da inferência estatística. No entanto, também, o modelo *Logit* tem em conta alguns pressupostos, tais como:

- Ausência de multicolinearidade entre as variáveis explicativas;
- Os erros do modelo seguem uma distribuição binomial e são independentes;
- A escala *Logit* é aditiva e linear;
- O contributo de cada variável explicativa é proporcional ao seu valor com um fator β_i ;
- O contributo das variáveis explicativas é constante e apresenta uma relação de independência com as outras variáveis explicativas.

Desta forma, Ohlson (1980) propôs um modelo de previsão de insolvência baseado num modelo *Logit*. Utilizou uma amostra composta por dados financeiros de 2 163 empresas (105 empresas insolventes e 2 058 empresas saudáveis) para o período temporal entre 1970 a 1976. O estudo recorreu à análise de 9 variáveis explicativas, 7 variáveis financeiras e 2 variáveis binárias, sendo estas:

1. $SIZE = \log \frac{Total\ Assets}{GNP\ price\ level\ index}$, variável utilizada como proxy da dimensão da empresa;
2. $TLTA = \frac{Total\ Liabilities}{Total\ Assets}$;
3. $WCTA = \frac{Working\ Capital}{Total\ Assets}$;
4. $CLCA = \frac{Current\ Liabilities}{Current\ Assets}$;
5. OENEG, 1 se passivo superior ao ativo, 0 caso contrário;
6. $NITA = \frac{Net\ Income}{Total\ Assets}$;
7. $FUTL = \frac{Cash\ Flow}{Total\ Liabilities}$;
8. INTWO, 1 se resultado líquido dos últimos dois anos foi negativo, 0 caso contrário;
9. $CHIN = \frac{Net\ Income_t - Net\ Income_{t-1}}{|Net\ Income_t| + |Net\ Income_{t-1}|}$.

Ohlson (1980) estimou três especificações *Logit* recorrendo às variáveis explicativas anteriores, de modo a aferir a probabilidade de falência: 1) um ano antes, 2) dois anos antes, e 3) entre o primeiro e segundo ano. Na Tabela 2 são apresentados os resultados obtidos nos 3 modelos estimados por Ohlson (1980), sendo de realçar que o autor conseguiu em todos os modelos uma capacidade preditiva acima dos 90%:

Tabela 2 – Resultados de Ohlson (1980)

Modelos	Capacidade Preditiva
1	0,9612
2	0,9555
3	0,9284

Fonte: Adaptado de Ohlson (1980).

Em termos de classificação de empresas como insolvente ou ativa, a inferência estatística

está sujeita a dois tipos de erros, que de acordo de Ohlson (1980) são definidos da seguinte forma: i) Erros do tipo I: quando se classifica uma empresa como insolvente e esta é saudável; ii) Erros do tipo II: quando se classifica uma empresa como saudável e esta é insolvente. O autor, inicialmente, considerou que os erros na categorização das empresas eram iguais, por essa razão, definiu um ponto de *cutoff* de 0,5. Assim, se o ponto de *cutoff* fosse superior a 0,5 a empresa era classificada como insolvente, em caso contrário a empresa era classificada como saudável. No entanto, o autor ao longo dos seus estudos procurou um valor que minimizasse a soma dos erros chegando a um ponto de *cutoff* de 0,038.

Os resultados obtidos por Ohlson (1980) mostram que existem quatro variáveis significativas que afetam a probabilidade de falência de uma empresa a um ano do acontecimento, sendo estes: a dimensão da empresa, a estrutura financeira, medidas de desempenho e a liquidez atual. O autor, considera que a capacidade preditiva do modelo é influenciada pelo *timing* com que a informação é obtida relativamente à data de insolvência das empresas e que a sua capacidade poderia aumentar com a utilização de outras variáveis explicativas. No entanto, apesar dos resultados obtidos serem ligeiramente inferiores aos obtidos pelos modelos que utilizaram a ADM, Ohlson (1980) considera que o uso do modelo *Logit* tem grandes vantagens quando comparado com a ADM, uma vez que é possível analisar amostras desproporcionais e não é necessária a verificação dos pressupostos exigidos na ADM, tal como referido anteriormente.

Depois de Ohlson (1980), vários autores aplicaram a mesma metodologia (modelos probabilísticos) nas suas investigações, entre os quais, Zavgren (1985) e Platt e Platt (1990). Em 1984, Zmijweski destacou-se por ter recorrido ao modelo *Probit* para realizar o seu estudo de previsão de insolvência. Este tipo de metodologia é semelhante à de Ohlson, sendo que a diferença reside no facto de na Análise *Probit* a probabilidade de insolvência acumulada ser dada pela função densidade da probabilidade associada à distribuição normal padronizada. O autor analisou a probabilidade de uma empresa entrar em falência através de três variáveis explicativas - liquidez, rentabilidade e alavancagem - recorrendo a uma amostra composta por 800 empresas insolventes e 40 empresas saudáveis, para o período de 1972 a 1978. Os resultados do estudo mostraram que o desempenho do modelo era semelhante ao obtido com o recurso a técnicas de amostragem aleatória.

2.2.2. Modelos específicos para PME

Falando agora mais especificamente acerca das pequenas e médias empresas, tendo em

conta que estas assumem um peso significativa em toda a economia mundial, muitos têm sido os estudos focados na previsão de insolvência deste tipo de empresas, uma vez que estas diferem das grandes empresas em termos de dimensão, volume de negócios e risco de crédito, necessitando assim de modelos individuais e específicos de previsão de falência.

O primeiro autor a investigar sobre este tipo de empresas foi Edmister, em 1972. Para a construção do modelo de previsão de insolvência, específico para PME, este recorreu à análise discriminante multivariada e a uma amostra de 562 empresas para uma análise a 1 ano e a outra amostra de 42 empresas para uma análise a 3 anos, para o período entre 1954 a 1969. Para a sua análise, o autor utilizou 19 rácios financeiros. No seu estudo, Edmister (1972) realça a importância da seleção dos rácios financeiros para se conseguir chegar a um modelo eficaz de previsão de falência, no entanto, não explica por que razão deve ser feita uma separação em termos de dimensão empresarial, ou seja, entre pequenas e grandes empresas.

Recentemente e seguindo o estudo de Edmister (1972), Altman e Sabato, em 2007, estudaram a previsão de falência, especificamente, para as PME americanas. Nesse estudo os autores desenvolveram um modelo de previsão de insolvência para um ano antes dessa situação ocorrer, analisando a eficácia do modelo em comparação com os modelos genéricos de previsão de insolvência criados até à data. O principal objetivo dos autores é analisar um conjunto de rácios financeiros ligados às PME dos Estados Unidos da América e identificar quais as variáveis que melhor explicavam o valor intrínseco de uma empresa. Para a sua investigação, recorreram a um modelo de regressão logística com dados em painel. A sua amostra era composta por dados financeiros de 2 010 PME americanas com vendas inferiores a 65 milhões de dólares, incluindo 120 empresas insolventes, para o período temporal de 1994 a 2002. Os autores selecionaram cinco categorias de rácios financeiros que melhor descrevem a situação financeira de uma empresa, sendo estes: liquidez, rentabilidade, endividamento, cobertura e atividade. Para cada categoria foram selecionados os rácios com maior poder explicativo, seguindo-se a aplicação do método *Forward Stepwise Selection*, de forma a obter o rácio com maior capacidade explicativa de cada categoria. Assim, selecionaram as seguintes variáveis:

$$1. \text{Liquidity} = \frac{\text{Cash}}{\text{Total Assets}};$$

$$2. \text{Profitability} = \frac{\text{EBITDA}}{\text{Total Assets}};$$

$$3. \text{Leverage} = \frac{\text{Short term debt}}{\text{Equity book value}};$$

$$4. \text{ Coverage} = \frac{\text{Retained earnings}}{\text{Total assets}};$$

$$5. \text{ Activity} = \frac{\text{EBITDA}}{\text{Interest expenses}}.$$

Com o objetivo de eliminar a dispersão dos valores dos rácios, que segundo os autores se pode dever a fatores como o setor de atividade ou a idade das empresas, estes efetuaram transformações logarítmicas nas suas variáveis. Na Tabela 3 são apresentados os coeficientes do modelo desenvolvido pelos autores.

Tabela 3 - Modelo de Altman e Sabato (2007)

Variável	Modelo sem transformações	Modelo com transformações
	logarítmicas	logarítmicas
<i>Cash/Total assets</i>	0,02	1,84
<i>EBITDA/Total assets</i>	0,18	4,09
<i>Short term debt/Equity book value</i>	-0,01	-1,13
<i>Retained earnings/Total assets</i>	0,08	4,32
<i>EBITDA/Interest expenses</i>	0,19	1,97

Fonte: Adaptado de Altman e Sabato (2007).

O estudo de Altman e Sabato (2007) teve duas fases distintas. Os resultados da primeira fase mostram que os sinais das variáveis são os esperados, sendo que todas as variáveis consideradas para a investigação mostraram ser estatisticamente significativas com capacidade preditiva de 75%. Na segunda fase, através da transformação logarítmica das variáveis, os autores conseguiram um incremento na capacidade de previsão de 75% para 87%. Os autores através destes dois tipos de análise mostraram que com um modelo específico para PME conseguiam um desempenho, em termos preditivos, superior ao modelo geral *Z-score*, que apresentou uma capacidade de previsão de 60%.

Em 2010, Altman, Sabato e Wilson, desenvolveram o trabalho de Altman e Sabato (2007), incluindo variáveis não financeiras para a construção de um modelo de previsão de insolvência. O objetivo do estudo foi melhorar a capacidade preditiva dos modelos de previsão de insolvência, através da combinação de variáveis financeiras com variáveis não financeiras, que, frequentemente, revelam informações importantes sobre a situação de uma empresa. Como a

generalidade das empresas de dimensão reduzida realizam contabilidade simplificada², este aspeto torna-se um ponto crítico para uma análise metódica e eficiente dos rácios financeiros relevantes na estimação de um modelo de previsão. Assim os autores basearam a sua investigação numa amostra de 5 749 188 empresas do Reino Unido (66 833 insolventes e 5 682 355 saudáveis) para o período temporal entre 2000 a 2007. Para a construção do modelo, os autores tiveram por base a especificação *Logit* utilizada em 2007 e construíram dois modelos distintos, um para empresas com informação financeira completa e outro para empresas com relatórios financeiros simplificados. Para além das variáveis financeiras, para a construção do modelo final foram utilizadas variáveis como a idade, a dimensão e o setor de atividade. Através deste estudo, os autores observaram um aumento de 13% na capacidade de previsão de insolvência a um ano, relativamente aos modelos anteriormente desenvolvidos. Por fim, os autores concluem que a construção de um modelo de previsão de insolvência específico para PME deve ter em conta não só variáveis financeiras, mas também informação não financeira, o que aumenta significativamente a precisão do modelo.

2.3 Objetivos do estudo

Como já referido, para o estudo sobre a insolvência deve ter-se em conta, para além da metodologia a adotar, as variáveis que serão utilizadas para estimar o modelo de previsão. Assim, como já relatado anteriormente, os estudos mais recentes sobre esta temática realçam a importância da inclusão de variáveis de natureza não financeira, como, por exemplo a idade, a dimensão, a localização geográfica, as habilitações académicas do gestor, entre outras, uma vez que os rácios financeiros podem ser facilmente manipulados por gestores, de forma a omitir a situação real da empresa, sendo estas variáveis, também, importantes para ter em conta os efeitos derivados da correlação com o ambiente macroeconómico em que a empresa está inserida.

O objetivo da presente investigação é a criação de um modelo de previsão de insolvência, para aplicação a PME do Norte e Centro de Portugal dos diversos setores de atividade. Para a especificação do modelo serão construídas variáveis com base na literatura já existente (p.e., Beaver, 1966; Altman, 1968; Ohlson, 1980), em conformidade com o tipo de estudo a realizar, ou seja, variáveis aplicáveis à realidade das PME de modo a construir um modelo logístico de previsão de insolvência, com recurso a variáveis de natureza contabilística e financeira.

² No regime simplificado os rendimentos devem ser inferiores a um determinado valor definido na lei e não precisa de contratar um contabilista certificado. No regime organizado, os rendimentos devem ser superiores a esse valor determinado legalmente e é obrigatório contratar um contabilista certificado.

3. Apresentação da Entidade Parceira

Como já referido anteriormente, o interesse da Comunidade Científica pelo estudo da previsão de insolvência deve-se à sua aplicação direta a contextos reais, sendo os modelos propostos utilizados pelos diversos agentes económicos da sociedade, como, por exemplo, o modelo proposto por Altman (1968). Os modelos de previsão de insolvência permitem fazer uma previsão sobre o futuro de uma empresa, auxiliar na tomada de decisão e na resolução de problemas a montante.

Assim sendo, a realização deste projeto de investigação deve-se ao facto de a nBanks, entidade parceira, ter como objetivo aplicar este estudo aos serviços que oferece aos seus clientes.

A nBanks é uma *fintech*³, do Norte de Portugal, relativamente recente, que tem as PME do Sul da Europa como principais alvos de mercado e, conseqüentemente, como principais clientes. A sua base de atuação assenta nos seguintes pontos:

- Agregação de contas bancárias e respetivos movimentos e informações subjacentes numa plataforma, sendo que essas mesmas informações são processadas e otimizadas, de forma a nBanks perceber o desempenho dos seus clientes tanto a nível financeiro, como económico e tributário, conseguindo assim, estabelecer padrões que permitem antecipar cenários;
- Pesquisa de produtos financeiros disponíveis no mercado (depósitos a prazo, soluções de investimento e de financiamento), de modo a perceber quais se enquadram mais com os seus clientes e, por conseguinte, estabelecimento de contacto entre a empresa e a entidade financeira e negociação do produto;
- Análise e avaliação de rácios como o EBIT (*Earnings before interest and taxes*) ou o EBITDA (*Earnings before interest, taxes, depreciation and amortization*), de forma às empresas terem noção do seu potencial;
- Facilitação da relação entre empresa e contabilista através de um serviço de “chat”, permitindo ao contabilista ter acesso a todos os dados que necessitar, em tempo real, como, por exemplo, contas bancárias, sendo apenas necessária a autorização do cliente;
- Sistema de comunicação com entidades financeiras, que permite que, caso um cliente queira iniciar um negócio com um banco, este não necessite de tanto tempo para tratar do processo, uma vez que a nBanks já tem acesso a vários indicadores de desempenho

³ Tipologia empresarial que combina tecnologia com oferta de serviços financeiros.

da empresa, permitindo assim, que se ultrapassem várias etapas e conduzindo a uma agilização do processo.

4. Metodologia

4.1. Descrição da Metodologia

De acordo com o objetivo deste estudo - propor um modelo de previsão de insolvência para aplicação a PME do Norte e Centro de Portugal - serão construídas variáveis de natureza contabilística e financeira e testadas através de diversas metodologias, tais como: a Análise Discriminante Multivariada, a Análise Logística e a Análise de *Matching*.

O primeiro passo na construção de um modelo de previsão de insolvência é a identificação das variáveis com maior capacidade explicativa da situação de insolvência. A amostra recolhida deve ser composta por empresas que entraram em situação de insolvência (e não recuperaram) em algum momento do horizonte temporal analisado e por empresas saudáveis, sendo que de preferência o grupo de empresas insolventes e de saudáveis devem ser consideradas como *peers* (setor de atividade, dimensão e localização semelhantes, bem como o seu nível de endividamento e de que forma esta variável pode influenciar a situação de insolvência).

Para a elaboração do modelo de previsão, ter-se-á como base a Análise Discriminante Multivariada (ADM) tal como em Altman (1968), e a regressão logística usada também por Ohlson (1980), sendo, posteriormente, aplicada a análise de *matching*, através da técnica de *Propensity Score Matching* (PSM) proposta por Rosenbaum and Rubin (1983).

Em relação à ADM, esta consiste numa técnica estatística que é utilizada para classificar qualquer observação num ou mais grupos *a priori* dependendo das características individuais de cada observação. É usada para classificar a variável dependente na sua forma qualitativa, ou seja, em insolvente ou não insolvente. Este método será utilizado como análise preliminar e complementar à que será a metodologia principal, a análise logística.

O modelo de *scoring*, designado por *Z-Score*, desenvolvido e apresentado por Altman em 1968, foi o primeiro modelo de previsão de insolvência conhecido baseado na ADM. Como já relatado, o objetivo do autor era explicar a previsão de falência através da utilização de vários rácios financeiros, analisados em simultâneo.

A ADM deve obedecer a um determinado conjunto de pressupostos, tais como: i) a distribuição normal das variáveis; ii) a homogeneidade das matrizes de variância-covariância; iii) a existência de diferenças significativas entre os grupos; e iv) a remoção de *outliers* (Ohlson, 1980). Em relação ao pressuposto da normalidade da distribuição das variáveis, são raras as evidências onde isso acontece e, por conseguinte, torna-se difícil ultrapassar essa limitação. Em relação à igualdade de matrizes de variância e covariância, esta deve ser observada uma vez que é uma

consequência da linearidade da função. A escolha das variáveis que apresentem diferenças médias significativas entre grupos (insolventes *versus* saudáveis) é importante e é um processo que pode ser realizado recorrendo, por exemplo, à análise univariada, e deve ser conduzida previamente já que vai influenciar os resultados obtidos. Por último, a presença de *outliers* deve ser previamente observada e tratada dado que em diversas vezes um teste estatístico pode não ser validado devido a observações anormais que deturpam o valor da média, sendo que a existência destas observações pode prejudicar a interpretação dos resultados resultantes de uma determinada amostra. O facto de a ADM não considerar dados não financeiros, como, por exemplo, a qualificação dos gestores, significa que é uma análise baseada unicamente em dados quantitativos e isto torna-se uma das limitações deste tipo de análise. Outra limitação deste tipo de análise é que, para se aplicar a análise discriminante é necessário saber qual o momento exato que antecede a falência, informação que nem sempre está disponível.

Em relação à regressão logística, esta é caracterizada pela definição de uma variável dependente (Y) dicotómica, i.e., uma *dummy* que neste contexto de previsão de insolvência identifica as empresas como “saudável” ou “insolvente”, ou seja, esta assume um carácter binário, sendo que as empresas serão classificadas como: i) insolventes no caso de $Y=1$, e ii) saudáveis no caso de $Y=0$.

O principal *outcome* dos modelos logísticos é a probabilidade de acontecimento de um dado evento, sendo essa probabilidade é influenciada por um conjunto de variáveis independentes medidas no momento anterior ao acontecimento, sendo que para este estudo foi considerado um período de desfasamento (i.e., 1 ano) das variáveis independentes face à ocorrência do acontecimento.

A utilização do modelo *Logit* tem como principal objetivo estimar a probabilidade de acontecimento de um evento em função das variáveis explicativas, sendo estimado pelo método da máxima verosimilhança. Este modelo apresenta vantagens em relação à ADM, como a possibilidade de utilização de informação qualitativa, a possibilidade de utilização de uma amostra desproporcional e a possibilidade de relaxamento da condição de normalidade na distribuição das variáveis.

Por fim, e tendo em conta a desproporcionalidade do número de observações e das características das empresas insolventes em relação às empresas saudáveis, aplicou-se a análise de *matching* de modo a ter uma amostra com um número de observação mais equilibrado entre grupos – empresas insolventes *versus* saudáveis - uma vez que é expectável obter uma baixa

capacidade preditiva em modelos que se baseiam em amostras desproporcionais. Para o efeito, e com o objetivo de ultrapassar este problema de enviesamento, aplicamos o método proposto por Rosenbaum e Rubin (1983) designado por *Propensity Score Matching* (PSM). Este método faz corresponder a cada observação de um grupo de tratamento – neste caso, uma empresa insolvente - uma observação de uma empresa do grupo de controlo – neste caso, uma empresa saudável - com características semelhantes, antes da situação de insolvência. O primeiro passo deste método baseia-se na estimação da probabilidade de uma empresa ser designada para o grupo de tratamento, dado um conjunto de variáveis explicativas da situação de insolvência, segundo um modelo probabilístico. Por conseguinte, o segundo passo consiste na análise de *matching*, onde cada observação tratada será comparada com a(s) empresa(s) não tratada(s) mais próxima(s), em termos médios e durante o período anterior ao período de tratamento.

5. Descrição da Amostra e das Variáveis

5.1. Amostra

Tendo em vista o objetivo do projeto, para esta investigação foram selecionadas empresas para as quais já foi decretada insolvência e empresas financeiramente “saudáveis” que possuam características similares, em termos de dimensão, setor e área geográfica. A amostra é composta por PME do Norte e Centro de Portugal, de acordo com a partição feita a partir das NUTS2⁴, e foi recolhida da base de dados ORBIS da Bureau van Dijk, entre 2011 e 2019.

Neste estudo foram consideradas como insolventes as empresas que apresentavam como status a seguinte designação *Active (insolvency proceedings)*, sendo esta informação obtida com recurso à base de dados ORBIS, sendo também disponibilizado pela base de dados informação mais detalhada sobre se a empresa estava insolvente ou saudável e a data de alteração desse status. As empresas consideradas como saudáveis apresentavam como status a designação *Active*.

Assim, uma empresa é considerada:

- *Active*: quando está a operar e se encontra numa situação financeira considerada como sustentável;
- *Active (insolvency proceedings)*: quando ainda continua em atividade, mas o processo de insolvência já foi desencadeado. Durante esse período, a empresa continua a operar, tentando pagar aos seus credores e retomar a sua atividade normal, assumindo-se a falta de pagamento aos credores como temporária e uma probabilidade moderada de regresso à sua atividade normal.

Numa primeira fase, a amostra recolhida respeitou os seguintes critérios:

1. Forma legal: Foram excluídas da amostra empresas públicas;
2. Dimensão: Micro, pequenas e médias empresas, com um mínimo de 5 trabalhadores;
3. Localização: Norte e Centro de Portugal, regiões com classificação de NUTS2;

⁴ O Decreto-Lei n.º 46/89 definiu os três níveis da Nomenclatura das Unidades Territoriais para Fins Estatísticos (NUTS) para as unidades territoriais portuguesas. NUTS II (NUTS 2) é constituído por sete unidades, cinco no continente e os territórios das Regiões Autónomas dos Açores e da Madeira.

4. Indústria: Empresas de todos os setores com exceção do setor financeiro e de indústrias altamente reguladas, como os setores da distribuição de energia e da distribuição de água e saneamento público.

Após a recolha respeitando os critérios acima referidos, a amostra obtida somava um total de 34 649 empresas do Norte de Portugal e 15 067 do Centro de Portugal. Relativamente ao status do Norte (Centro) contavam-se 626 (263) empresas insolventes e 34 023 (14 804) empresas saudáveis.

Por conseguinte, após a recolha dos dados foram, posteriormente, aplicados os seguintes filtros:

- N° de trabalhadores: >10 e <250;
- Cada empresa tem de ter um mínimo de 3 anos consecutivos de observações;
- A informação contabilística necessária ao cálculo de todas as variáveis do modelo tem de estar disponível nos dois anos anteriores ao ano de insolvência.

A amostra final obtida é bastante desproporcional face ao número de observações de empresas insolventes e de empresas saudáveis. A amostra é constituída por um total de 18 622 empresas do Norte de Portugal e 7 633 do Centro de Portugal. Relativamente ao status do Norte (Centro) contavam-se 312 (128) empresas insolventes e 18 310 (7 505) empresas saudáveis.

5.2. Variáveis

5.2.1. Variável Dependente

Para esta investigação, a variável dependente será a classificação da empresa como insolvente ou saudável. Desta forma, torna-se fundamental definir de forma rigorosa o conceito de insolvência no âmbito do presente trabalho. Deste modo, serão consideradas como insolventes, as empresas que apresentam o seguinte status: *Active (insolvency proceedings)*. Deste modo, a variável dependente é designada por *Status*. Esta informação tem em conta o registo que se encontra disponível na base de dados ORBIS.

5.2.2. Variáveis Explicativas

Tendo como base a literatura mais relevante em relação ao tema desta investigação (p.e,

Beaver, 1966; Altman, 1968; Ohlson, 1980; Altman e Sabato, 2007; Altman et al., 2010), para a construção do modelo de previsão irão ser utilizados um conjunto de indicadores económico-financeiros apresentados de seguida na Tabela 4. As variáveis serão testadas numa análise preliminar com vista a avaliar a sua capacidade preditiva e, por conseguinte, a sua eventual inclusão como variáveis explicativas determinantes da insolvência, tal como explicado no Subcapítulo dedicado à Metodologia. Deve-se, também, realçar que todas as variáveis contínuas serão sujeitas ao procedimento de *winsorizing* em 1% em cada cauda, de forma a mitigar problemas com potenciais *outliers* na análise empírica.

Tabela 4 – Variáveis contruídas para o estudo

Denominação	Variável	Rácio
Liquidez Geral	X1	Ativo Corrente / Passivo Corrente
Liquidez Imediata	X2	Caixa e Equivalentes / Passivo Corrente
Rácio de Fundo de Maneio	X3	Fundo de Maneio / Ativo Total
Estrutura de Endividamento	X4	Passivo Corrente / Passivo Total
Rentabilidade Económica	X5	EBIT / Ativo Total
Rentabilidade do Ativo	X6	Resultado Líquido / Ativo Total
Margem EBIT	X7	EBIT / Vendas
Rentabilidade das Vendas	X8	Resultado Líquido / Vendas
Rácio Passivo – Capital Próprio	X9	Passivo Total / Capital Próprio
Rácio de Endividamento	X10	Passivo Total / Ativo Total
Solvabilidade	X11	Capital Próprio / Passivo Total
Rácio de Cobertura de Juros	X12	EBITDA / Juros Pagos
Rotação do Ativo	X13	Vendas / Ativo Total
Rácio Capital Próprio – Vendas	X14	Capital Próprio / Vendas
Dimensão	X15	Log (Ativo Total)
Idade	X16	Ano Corrente – Ano da Constituição

A tabela 4 apresenta as variáveis construídas na presente investigação.

De seguida, apresenta-se a descrição detalhada de cada um dos indicadores apresentados na Tabela 4. É ainda importante salientar a natureza genérica de cada comentário, dado que a análise de rácios contabilísticos deve ser enquadrada de acordo com o setor de atividade da empresa.

- **Liquidez Geral:** $\text{Ativo Corrente} / \text{Passivo Corrente}$ – Fornece informações sobre a capacidade de a empresa pagar os seus compromissos de curto prazo. Sempre que possível deverá ser superior a 1, o que a não acontecer exige o recurso aos capitais não correntes para financiar esse défice. Assume-se que valores inferiores a 1 podem revelar dificuldades de tesouraria no curto prazo, que, por conseguinte, aumenta a probabilidade de insolvência de uma empresa (Jackendoff, 1962).
- **Liquidez Imediata:** $\text{Caixa e Equivalentes} / \text{Passivo Corrente}$ – Mede a quantidade de ativos líquidos em relação ao passivo circulante de uma empresa. Ativos líquidos são ativos circulantes que podem ser rapidamente convertidos em dinheiro, enquanto os passivos circulantes são dívidas ou obrigações de uma empresa que devem ser pagas a curto prazo. Empresas com margens mais reduzidas de liquidez, em princípio, estão mais propensas a uma situação de insolvência. (Hopp & Leite, 1989).
- **Rácio de Fundo de Maneio:** $\text{Fundo de Maneio} / \text{Ativo Total}$ – Este rácio dá-nos a diferença entre o valor do Ativo Corrente e Passivo Corrente em relação ao Ativo Total. Rácios com valores baixos estão, geralmente, associados a um risco de insolvência mais elevado (Jackendoff, 1962).
- **Estrutura de Endividamento:** $\text{Passivo Corrente} / \text{Passivo Total}$ – Este rácio identifica, como o próprio nome indica, a estrutura de endividamento de uma empresa, ou seja, a distribuição da exigibilidade das obrigações da empresa ao longo do tempo. Em geral, se uma empresa apresentar um passivo de curto prazo superior ao passivo de longo prazo, isso pode significar que a empresa precisa de financiar constantemente o seu ciclo de negócio, o que pode indiciar maior vulnerabilidade perante os credores de curto prazo, aumentando assim a probabilidade de insolvência (Rodrigues, 2018).
- **Rentabilidade Económica:** $\text{EBIT} / \text{Ativo Total}$ – Este rácio compara os resultados operacionais, antes de juros e impostos, de uma empresa com o seu Ativo Total. Quanto mais elevado for este rácio melhor será o desempenho operacional da empresa, uma vez que significa que os ativos da empresa estão a ser bem utilizados e a produzir bons resultados, já a situação contrária pode ser um indiciador de uma situação que conduza a um futuro estado de insolvência, principalmente quando o mau desempenho se torna cumulativo ao longo dos anos (Beaver, 1966).
- **Rentabilidade do Ativo:** $\text{Resultado Líquido} / \text{Ativo Total}$ – Este rácio fornece informação sobre a capacidade dos ativos de uma empresa gerarem rendimento. Naturalmente,

quanto mais elevados os valores deste rácio, melhor será considerado o desempenho da empresa, ou seja, significa que os ativos da empresa estão a ser explorados de forma rentável (Moreira & Tavares, 2014).

- Margem EBIT: $\text{EBIT} / \text{Vendas}$ – Este rácio caracteriza a Rentabilidade Operacional das Vendas, ou seja, apresenta a contribuição na margem operacional, antes de juros e impostos, de cada unidade monetária gerada pelas vendas de uma empresa. Valores considerados baixos ou negativos em anos consecutivos poderão indiciar uma potencial situação de insolvência (Altman & Sabato, 2007).
- Rentabilidade das Vendas: $\text{Resultado Líquido} / \text{Vendas}$ – A Rentabilidade das Vendas permite avaliar a eficiência operacional de uma empresa. Ou seja, este rácio permite conhecer a margem líquida que uma empresa gera por cada unidade monetária obtida pelas suas vendas. Valores elevados deste rácio ao longo do tempo significam que a empresa está a tornar-se cada vez mais eficiente. Por outro lado, valores baixos acumulados durante anos de atividade podem significar problemas financeiros e uma maior propensão para a insolvência (Jagels & Colman, 2004).
- Rácio Passivo – Capital Próprio: $\text{Passivo Total} / \text{Capital Próprio}$ – Este rácio relaciona o Passivo Total e o Capital Próprio de uma empresa, ou seja, a relação entre as suas obrigações e o seu património. Quando este rácio apresenta valores elevados isso pode significar que a atividade da empresa se baseia em financiamentos sob a forma de crédito, o que a longo prazo poderá contribuir para uma situação de insolvência (Pereira, Tavares & Pacheco, 2015).
- Rácio de Endividamento: $\text{Passivo Total} / \text{Ativo Total}$ – Este rácio traduz a percentagem do Ativo que foi financiado pelo Passivo. Quanto mais elevado este rácio, maior o risco de insolvência associado (Ohlson, 1980).
- Solvabilidade: $\text{Capital Próprio} / \text{Passivo Total}$ – Permite analisar a capacidade da empresa cumprir as suas obrigações a médio e longo prazo. Valores baixos correspondem a empresas com um nível endividamento elevado e com maior probabilidade de insolvência, ou seja, quanto mais elevado este rácio, maior a estabilidade financeira da empresa (Altman, 1968).
- Rácio de Cobertura de Juros: $\text{EBITDA} / \text{Juros Pagos}$ – É utilizado para determinar a facilidade com que uma empresa consegue pagar as suas despesas com juros relacionadas com os compromissos financeiros. Valores abaixo de 1 significam que a

empresa não consegue gerar ganhos para pagar as despesas de financiamento que permitem que esta realize a sua atividade, o que no longo prazo pode levar a uma situação de insolvência dado que a empresa não se conseguirá financiar (Cherobim, Junior & Rigo, 2010).

- Rotação do Ativo: $Vendas / Ativo\ Total$ – Este rácio mede o grau de eficácia do Ativo através das vendas e a frequência com que o valor do Ativo se renova em função do valor das vendas. Indica, também, o número de unidades monetárias vendidas por cada unidade monetária investida pela empresa. Segundo Altman, (1968), quanto maior for este rácio, maior será, em princípio, a eficiência da empresa e menor o risco de insolvência.
- Rácio Capital Próprio – Vendas: $Capital\ Próprio / Vendas$ – Este rácio ilustra a relação entre o Capital Próprio e as Vendas, ou seja, traduz o património líquido que uma empresa consegue gerar através das suas vendas. Valores elevados significam que por cada unidade vendida a empresa tem uma elevada margem de lucro, o que, a longo prazo, lhe permitirá aumentar o seu património; caso contrário, a degradação deste rácio ao longo do tempo pode indicar uma situação de insolvência eminente (Fernandes, Peginho & Vieira, 2012).
- Dimensão: Logaritmo do Ativo Total – A dimensão da empresa será captada pelo logaritmo do total dos ativos; funciona como uma *proxy* da dimensão da empresa (Domenichelli, 2018).
- Idade: (Logaritmo (Ano corrente – Ano da Constituição)) – Esta variável foi incluída para eliminar problemas de heteroscedasticidade (Domenichelli, 2018), relacionados com a escala das variáveis, este indicador será medido pela diferença entre o ano em causa e o ano da criação da empresa.

5.3. Estatísticas Descritivas

As Tabelas 5, 6 e 7 apresentam as estatísticas descritivas das variáveis explicativas, para o período de 2011 a 2019. A Tabela 5 apresenta as estatísticas descritivas para a amostra total, enquanto a Tabela 6 apresenta as estatísticas descritivas para as subamostras compostas por empresas do Norte e Centro de Portugal, representadas por NUTS2.1 e NUTS2.2, respetivamente. Já a Tabela 7 apresenta as estatísticas descritivas por *Status*, ou seja, faz a distinção entre empresas saudáveis (0) e empresas insolventes (1).

Tabela 5 – Estatísticas Descritivas: Amostra Total

Variável	Nº Obs.	Média	P ₂₅	P ₅₀	P ₇₅	DP
Liquidez Geral	108 366	2,18	1,14	1,73	2,65	1,74
Liquidez Imediata	108 193	0,47	0,06	0,20	0,55	0,74
Rácio de Fundo de Maneio	144 192	0,23	0,04	0,22	0,40	0,26
Estrutura de Endividamento	108 366	0,62	0,43	0,62	0,84	0,25
Rentabilidade Económica	144 792	0,04	0,01	0,04	0,09	0,13
Rentabilidade do Ativo	144 792	0,02	0,00	0,02	0,06	0,12
Margem EBIT	144 792	0,03	0,01	0,03	0,07	0,10
Rentabilidade das Vendas	144 792	0,01	0,00	0,02	0,05	0,10
Rácio Passivo – Capital Próprio	144 792	2,90	0,59	1,50	3,25	8,13
Rácio de Endividamento	144 792	0,67	0,45	0,65	0,82	0,36
Solvabilidade	144 758	0,65	0,22	0,53	1,20	1,69
Rácio de Cobertura de Juros	115 134	0,00	3,73	11,75	45,52	0,00
Rotação do Ativo	144 792	1,61	0,84	1,28	1,96	1,25
Rácio Capital Próprio – Vendas	144 792	0,36	0,11	0,26	0,49	0,47
Dimensão	144 792	6,52	5,74	6,58	7,35	1,16
Idade	63 547	19,25	9,00	17,00	26,00	13,70

A Tabela 5 apresenta as seguintes estatísticas descritivas: Número de Observações, Média, Percentil 25.º, Percentil 50.º que corresponde à Mediana, Percentil 75.º e Desvio Padrão (D.P.). As variáveis são as seguintes: Liquidez Geral, Liquidez Imediata, Rácio de Fundo de Maneio, Estrutura de Endividamento, Rentabilidade Económica, Rentabilidade do Ativo, Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio, Rácio de Endividamento, Solvabilidade, Rácio de Cobertura de Juros, Rotação do Ativo, Rácio Capital Próprio-Vendas, Dimensão e Idade, construídas conforme descrito no Subcapítulo 5.2.2.

Como se pode observar na Tabela 5, em média, o Ativo Corrente é 2,18 vezes superior ao Passivo Corrente (Liquidez Geral). O fundo de maneio (diferença entre o Ativo Corrente e o Passivo Corrente) representa, em média, 23% do Ativo Total (Rácio de Fundo Maneio). A Rotação do Ativo é, em média, 1,61 e a Solvabilidade é de 0,65. O Passivo representa, em média, 67% do Ativo (Rácio de Endividamento).

A Tabela 6 reporta que os valores médios das variáveis Liquidez Imediata, Rácio de Fundo de Maneio, Estrutura de Endividamento, Rácio Passivo-Capital Próprio, Rácio de Endividamento e Rotação do Ativo são superiores nas empresas do Norte relativamente às empresas do Centro, e inferiores no que respeita a Liquidez Geral, Rentabilidade Económica, Margem EBIT, Rentabilidade das Vendas, Solvabilidade, Rácio Capital Próprio-Vendas, Dimensão e Idade.

Tabela 6 – Estatísticas Descritivas por NUTS2

Variável	NUTS2.1 Norte				NUTS2.2 Centro			
	Nº Obs.	Média	P ₅₀	DP	Nº Obs.	Média	P ₅₀	DP
Liquidez Geral	61 017	2,15	1,72	1,69	26 277	2,24	1,78	1,79
Liquidez Imediata	60 936	0,47	0,21	0,72	26 231	0,44	0,16	0,73
Rácio de Fundo de Maneio	81 741	0,25	0,24	0,26	34 742	0,21	0,20	0,25
Estrutura de Endividamento	61 017	0,64	0,65	0,25	26 277	0,58	0,57	0,25
Rentabilidade Económica	82 145	0,04	0,04	0,13	34 817	0,05	0,04	0,12
Rentabilidade do Ativo	82 145	0,02	0,02	0,12	34 817	0,02	0,02	0,11
Margem EBIT	82 145	0,03	0,03	0,10	34 817	0,04	0,04	0,10
Rentabilidade das Vendas	82 145	0,01	0,02	0,09	34 817	0,02	0,02	0,10
Rácio Passivo – Capital Próprio	82 145	2,96	1,58	8,08	34 817	2,75	1,42	7,79
Rácio de Endividamento	82 145	0,68	0,67	0,36	34 817	0,63	0,63	0,33
Solvabilidade	82 127	0,63	0,50	1,59	34 810	0,71	0,59	1,81
Rácio de Cobertura de Juros	64 168	0,00	11,17	0,00	27 821	0,00	10,70	0,00
Rotação do Ativo	82 145	1,64	1,33	1,21	34 817	1,48	1,15	1,23
Rácio Capital Próprio – Vendas	82 145	0,34	0,24	0,44	34 817	0,43	0,31	0,52
Dimensão	82 145	6,44	6,50	1,16	34 817	6,80	6,86	1,07
Idade	35 875	14,15	15,96	2,36	15 431	16,28	18,92	2,20

A Tabela 6 apresenta as estatísticas descritivas por NUTS: Número de Observações, Média, Percentil 50.º que corresponde à Mediana e Desvio Padrão (D.P.). As variáveis são as descritas no Subcapítulo 5.2.2.

Observa-se na Tabela 7 que os valores médios das variáveis Estrutura de Endividamento, Rácio de Endividamento, Solvabilidade e Rotação do Ativo são superiores nas empresas saudáveis relativamente às empresas insolventes, e inferiores no que respeita a Liquidez Geral, Liquidez Imediata, Rácio de Fundo de Maneio, Rentabilidade do Ativo, Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio, Rácio Capital Próprio-Vendas, Dimensão e Idade.

Tabela 7 – Estatísticas Descritivas por *Status*

Variável	<i>Status</i> 0				<i>Status</i> 1			
	Nº Obs.	Empresas saudáveis			Nº Obs.	Empresas insolventes		
Média		P ₅₀	D P	Média		P ₅₀	DP	
Liquidez Geral	85 823	2,18	1,73	1,72	1 471	2,35	1,85	1,85
Liquidez Imediata	85 698	0,46	0,19	0,72	1 469	0,50	0,21	0,80
Rácio de Fundo de Maneio	114 545	0,23	0,22	0,26	1 938	0,24	0,25	0,26
Estrutura de Endividamento	85 823	0,62	0,62	0,25	1 471	0,61	0,62	0,25
Rentabilidade Económica	115 017	0,05	0,04	0,13	1 945	0,05	0,04	0,12
Rentabilidade do Ativo	115 017	0,02	0,02	0,12	1 945	0,03	0,02	0,11
Margem EBIT	115 017	0,03	0,03	0,10	1 945	0,04	0,03	0,10
Rentabilidade das Vendas	115 017	0,01	0,02	0,09	1 945	0,02	0,02	0,09
Rácio Passivo – Capital Próprio	115 017	2,89	1,53	7,98	1 945	3,52	1,61	8,90
Rácio de Endividamento	115 017	0,67	0,66	0,35	1 945	0,66	0,65	0,34
Solvabilidade	114 992	0,67	0,53	1,66	1 945	0,65	0,54	1,63
Rácio de Cobertura de Juros	90 443	0,00	11,02	0,00	1 546	0,00	11,90	0,00
Rotação do Ativo	115 017	1,59	1,28	1,22	1 945	1,55	1,24	1,25
Rácio Capital Próprio – Vendas	115 017	0,37	0,26	0,47	1 945	0,41	0,28	0,52
Dimensão	115 017	6,55	6,61	1,14	1 945	6,67	6,69	1,19
Idade	50 467	14,73	16,95	2,32	839	16,28	20,01	2,59

A Tabela 7 apresenta as estatísticas descritivas por *Status*: Número de Observações, Média, Percentil 50.º que corresponde à Mediana e Desvio Padrão (D.P.). As variáveis são as descritas no Subcapítulo 5.2.2.

Nas Tabelas 8 e 9 são apresentadas, respetivamente, as diferentes descrições da amostra ao longo do período temporal da amostra em função do *Status* (insolvente *vs* saudável) e por tipo de indústria, cuja definição encontra-se na Tabela A do Apêndice, em função da região.

A informação na Tabela 8 traduz uma maior representatividade de empresas nos anos 2018 e 2019. E como esperado, a Tabela 9 mostra que a maioria das observações estão concentradas na Zona Norte.

Tabela 8 – Descrição da amostra por ano e *Status*

Ano	<i>Status</i> = 0 Empresas saudáveis Nº Obs.	<i>Status</i> = 1 Empresas insolventes Nº Obs.	Total
2011	12 620	203	12 823
2012	12 677	204	12 881
2013	13 088	212	13 300
2014	14 030	230	14 260
2015	15 462	257	15 719
2016	16 819	289	17 108
2017	17 863	322	18 185
2018	19 080	328	19 408
2019	19 231	343	19 574
Total	140 870	2 388	143 258

A Tabela 8 apresenta uma descrição da amostra por ano e por *Status*, tendo em conta o número de observações totais.

Tabela 9 - Descrição da amostra por tipo de indústria e por região

Indústria	Norte Nº Obs.	Centro Nº Obs.	Total
3	44 491	17 820	62 311
5	325	194	519
6	14 278	0	14 278
7	18 720	9 946	28 666
8	3 503	2 907	6 410
9	7 600	4 441	12 041
10	1 211	589	1 800
13	3 039	1 530	4 569
14	3 190	1 645	4 835
16	1 327	723	2 050
17	2 115	2 179	4 294
18	385	231	616
19	607	262	869
Total	100 791	42 467	143 258

A Tabela 9 apresenta uma descrição da amostra por tipo de indústria e por região, tendo em conta o número de observações totais. Os diferentes tipos de indústria estão definidos na Tabela A do Apêndice.

A Tabela 10 apresenta a descrição da amostra por região (NUTS2) e por *Status* (insolvente

vs saudável).

Tabela 10 – Descrição da amostra por região e por *Status*

NUTS2	<i>Status</i> = 0 Empresas saudáveis Nº Obs.	<i>Status</i> = 1 Empresas insolventes Nº Obs.	Total
Norte	99 109	1 682	100 791
Centro	41 761	706	42 467
Total	140 870	2 388	143 258

A Tabela 10 apresenta uma descrição da amostra por região e por *Status*, tendo em conta o número de observações totais.

5.4. Análise Univariada

Foi através da análise univariada que se realizaram as primeiras investigações sobre modelos de previsão de insolvência. A investigação empírica de Beaver (1966) destaca-se ainda nos dias de hoje, como um dos estudos mais importantes, com base na análise univariada de variáveis de previsão de insolvência. No entanto, o próprio Beaver (1966) reconhece que a capacidade preditiva de um modelo de previsão pode melhorar se os rácios forem analisados em simultâneo, através da ADM, e não de forma isolada.

Com esta análise pretende-se testar as diferenças às médias e às medianas entre as empresas insolventes e as empresas saudáveis. Assim sendo, realizou-se um teste paramétrico às médias – *t-test* - e um teste não paramétrico às medianas – Wilcoxon-Mann-Whitney *rank sum test* - através da comparação das médias e das medianas das variáveis quantitativas dos dois grupos independentes. Primeiro em função das NUTS2 (Norte e Centro) e depois em função do *Status*, insolventes vs saudáveis. A Tabela 11 apresenta os resultados.

Após a análise à Tabela 11, observa-se que, à exceção das variáveis Rentabilidade Económica, Rentabilidade do Ativo e Rácio de Cobertura de Juros, a diferença entre as médias e as medianas das restantes variáveis apresentam significância estatística ao nível de 1% no que respeita à análise em função das NUTS2 (Norte e Centro).

Tabela 11 – Teste às médias e às medianas em função das NUTS2 e do *Status*

Variável	NUTS		<i>Status</i>	
	Diferenças entre médias	Diferenças entre medianas	Diferenças entre médias	Diferenças entre medianas
	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>	<i>p-value</i>
Liquidez Geral	0,0000	0,0000	0,0004	0,0003
Liquidez Imediata	0,0000	0,0000	0,0622	0,0669
Rácio de Fundo de Maneio	0,0000	0,0000	0,2458	0,1735
Estrutura de Endividamento	0,0000	0,0000	0,0311	0,0345
Rentabilidade Económica	0,0012	0,9336	0,0618	0,3297
Rentabilidade do Ativo	0,0002	0,0229	0,0588	0,4143
Margem EBIT	0,0000	0,0000	0,0147	0,2339
Rentabilidade das Vendas	0,0000	0,0000	0,0338	0,2507
Rácio Passivo – Capital Próprio	0,0000	0,0000	0,0017	0,2281
Rácio de Endividamento	0,0000	0,0000	0,3173	0,8530
Solvabilidade	0,0000	0,0000	0,6183	0,8399
Rácio de Cobertura de Juros	0,8423	0,0097	0,6393	0,2681
Rotação do Ativo	0,0000	0,0000	0,1681	0,0210
Rácio Capital Próprio – Vendas	0,0000	0,0000	0,0001	0,0157
Dimensão	0,0000	0,0000	0,0000	0,0000
Idade	0,0000	0,0000	0,0038	0,0000

A Tabela 11 apresenta o teste às diferenças entre médias e medianas em função das NUTS2 e do *Status*. As variáveis são as descritas no Subcapítulo 5.2.2.

Relativamente à análise em função do *Status* das empresas, apenas nas variáveis Liquidez Geral e Dimensão, a diferença entre as médias e as medianas é estatisticamente significativa ao nível de 1%.

Após comparação da informação da Tabela 6 – Estatísticas Descritivas por NUTS2 – com a informação disponível na Tabela 11 observa-se que o rácio Rotação do Ativo é superior nas empresas do Norte. Assim, conclui-se que as empresas do Norte são mais eficientes em relação às do Centro, o que por conseguinte, conduz à conclusão de que as empresas do Norte apresentam menor risco de insolvência. Por outro lado, as empresas do Centro apresentam valores inferiores, em relação às empresas do Norte, no Rácio de Endividamento e Solvabilidade, o que sugere que, em termos médios, que as empresas do Norte são ligeiramente mais propensas à situação de insolvência.

Em relação à comparação entre o grupo de empresas insolventes e o grupo de empresas saudáveis, os grupos apresentam diferenças nas médias estatisticamente significativas ao nível

de, pelo menos, 10%, exceto Rácio de Fundo de Maneio, Rácio de Endividamento e Solvabilidade, Rácio de Cobertura de Juros e Rotação do Ativo, sendo que no caso das diferenças entre medianas, somente o Rácio de Liquidez Geral, Liquidez Imediata, Estrutura de Endividamento, Rácio Capital Próprio – Vendas, Rotação do Ativo, Dimensão e Idade.

6. Análise Multivariada

6.1. Análise Discriminante Múltipla

A primeira metodologia a ser aplicada na Análise Multivariada foi a Análise Discriminante Múltipla (ADM), que é uma das técnicas mais utilizadas no estudo dos modelos de previsão de insolvência.

A ADM tem como principal objetivo complementar a análise logística no sentido de suportar a decisão de seleção das variáveis que irão ser incluídas no modelo final de previsão de falência.

Assim, esta análise encontra-se estruturada da seguinte forma⁵:

- classificação das empresas;
- estimação da função discriminante;
- análise da matriz de correlação de Pearson;
- análise da variância designada por ANOVA;
- por fim, seleção das variáveis.

O primeiro passo da ADM centra-se na classificação das empresas como insolventes com o valor de um (1), enquanto as empresas saudáveis são classificadas com o valor de zero (0). Esta classificação é feita de acordo com as variáveis explicativas já apresentadas e analisadas no Subcapítulo 5.2.2.

Deste primeiro passo da análise discriminante resulta uma tabela denominada por *Confusion Table*, que classifica cada empresa como insolvente ou saudável, de acordo com as variáveis explicativas que a ADM considera como características diferenciadoras de cada grupo (insolvente/saudável).

⁵ Ver, p.e.: SAGE, 2015, p. 590-604.

Tabela 12 – Classificação das empresas por *Status*

<i>True Status</i>	Classificação		Total
	Empresas saudáveis 0	Empresas insolventes 1	
0	18 035 57,00%	13 607 43,00%	31 642 100,00%
1	242 42,76%	324 57,24%	566 100,00%
Total	18 277 56,75%	13 931 43,25%	32 208 100,00%

A Tabela 12 classifica as empresas consideradas insolventes com o valor um (1) e as empresas consideradas como saudáveis com o valor zero (0), de acordo com as variáveis explicativas utilizadas no presente estudo, descritas no Subcapítulo 5.2.2.

Com base na classificação das empresas como insolventes ou saudáveis, observamos que 57,00% das empresas classificadas como saudáveis estão corretamente classificadas e que 43,00% das empresas classificadas como saudáveis deveriam ser consideradas insolventes, de acordo com as variáveis explicativas utilizadas para este estudo. Estamos perante erros do tipo I e II, conforme a classificação proposta por Ohlson (1980) e apresentada na Tabela 13. O Erro do Tipo I, resulta da classificação de empresas saudáveis como insolventes, já o Erro do Tipo II, resulta da classificação de empresas insolventes como saudáveis.

Tabela 13 – Erros tipos I e II

Previsão do Resultado	Verdadeiro Positivo (y=1) Empresas Insolventes	Falso Positivo Erro tipo I: empresas saudáveis classificadas como insolventes (1 – Specificity)
	Falso Negativo Erro tipo II: empresas insolventes classificadas como saudáveis (1 – Sensitivity)	Verdadeiro Negativo (y=0) Empresas Saudáveis

A Tabela 13 permite perceber se as empresas estão corretamente classificadas, cruzando as empresas observadas com as empresas classificadas pela aplicação da ADM, constituindo um instrumento de análise da capacidade discriminante do modelo. A partir desta Tabela é possível conhecer os erros do tipo I e II.

O próximo passo da ADM é a seleção das variáveis que permitem diminuir a probabilidade de erro na estimação do modelo de previsão e que demonstram uma capacidade explicativa capaz de identificar uma situação de insolvência iminente.

O peso de cada variável na classificação realizada no primeiro passo da ADM de empresas insolventes/saudáveis, independentemente do grupo, é representado pelos seus coeficientes *unstandardized*, sendo por isso considerados na função discriminante.

A análise dos coeficientes *unstandardized* está limitada à discriminação das observações entre grupos e à escala das variáveis; por exemplo, a Dimensão é medida como o logaritmo dos ativos, indicador que será, em média, superior aos rácios de, p.e., Liquidez, logo a análise *unstandardized* atribuirá maior peso discriminativo à variável Dimensão. Deste modo, para eliminar problemas de escala entre variáveis foi realizado o processo de estandardização dos coeficientes. Assim, os coeficientes *standardized* discriminam a capacidade explicativa da variável dentro de cada grupo, o que permite observar a contribuição de cada variável para explicar a situação de insolvência. Assim, na análise *standardized* o peso de cada variável deve ser interpretado em função do seu valor absoluto (i.e., independentemente do seu sinal).

Tabela 14 – Coeficientes da Função Discriminante

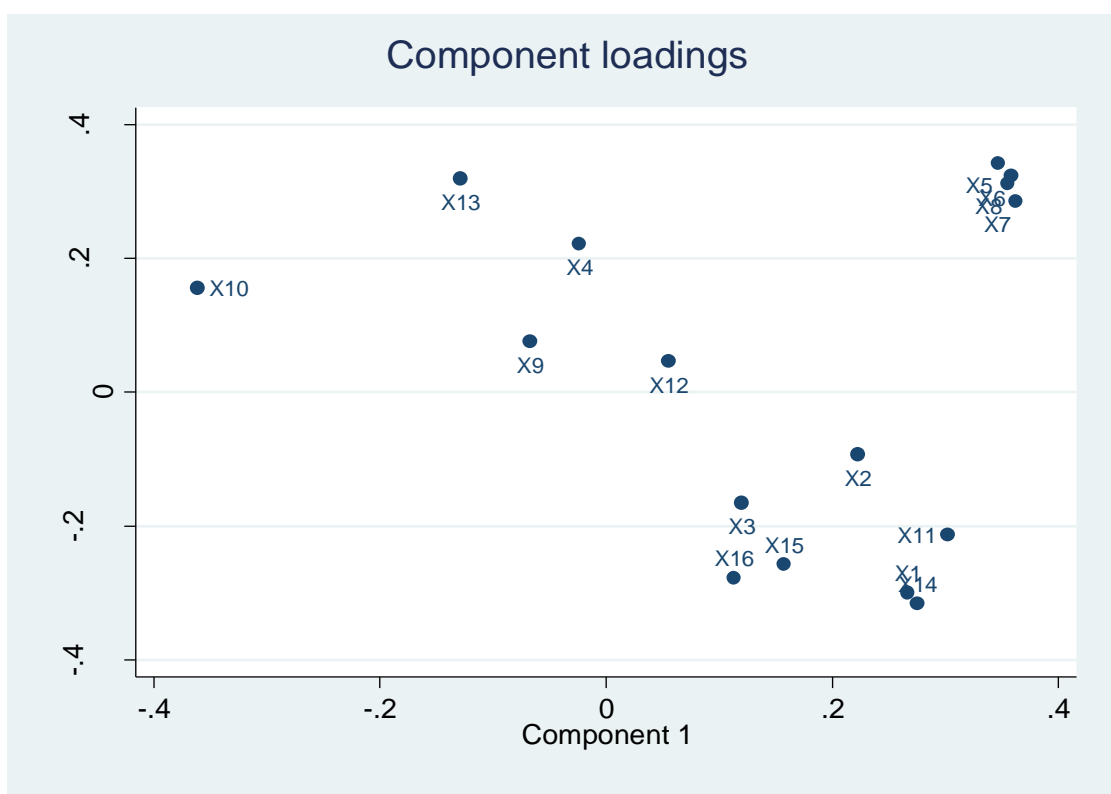
Variável	Coeficientes <i>Unstandardized</i>	Coeficientes <i>Standardized</i>
Liquidez Geral	-0,050	-0,083
Liquidez Imediata	0,122	0,083
Rácio de Fundo de Maneio	1,053	0,252
Estrutura de Endividamento	-1,401	-0,334
Rentabilidade Económica	-4,696	-0,506
Rentabilidade do Ativo	7,311	0,731
Margem EBIT	-14,846	-1,348
Rentabilidade das Vendas	9,272	0,802
Rácio Passivo – Capital Próprio	-0,065	-0,504
Rácio de Endividamento	-0,080	-0,024
Solvabilidade	-0,165	-0,187
Rácio de Cobertura de Juros	0,000	0,222
Rotação do Ativo	-0,343	-0,362
Rácio Capital Próprio – Vendas	0,235	0,101
Dimensão	-0,227	-0,249
Idade	-0,749	-0,601
Constante	5,405	-

A Tabela 14 apresenta os coeficientes *unstandardized* e os coeficientes *standardized*, que correspondem à 2.ª etapa da ADM. As variáveis são as descritas no Subcapítulo 5.2.2.

A transformação dos coeficientes *unstandardized* em *standardized* permite observar as variáveis com maior poder explicativo da situação de insolvência de uma empresa. Deste modo é possível selecionar as variáveis que apresentam os coeficientes mais elevados e que serão consideradas nas próximas etapas da análise multivariada.

Assumindo como pressuposto da ADM que as variáveis com maior poder explicativo na previsão de insolvência são as que possuem os coeficientes mais elevados, independentemente do sinal, as variáveis são ordenadas em função do seu peso, sendo de seguida efetuada uma observação da representação gráfica dos coeficientes para verificar a existência de variáveis sobrepostas. Esta técnica serve de apoio à decisão de seleção de variáveis.

Gráfico 1 - Coeficientes *standardized*



O Gráfico 1 apresenta o peso de cada variável em função dos coeficientes *standardized*. As variáveis são as descritas no Subcapítulo 5.2.2. e são as seguintes: X1 - Liquidez Geral, X2 - Liquidez Imediata, X3 - Rácio de Fundo de Maneio, X4 - Estrutura de Endividamento, X5 - Rentabilidade Económica, X6 - Rentabilidade do Ativo, X7 - Margem EBIT, X8 - Rentabilidade das Vendas, X9 - Rácio Passivo-Capital Próprio, X10 - Rácio de Endividamento, X11 - Solvabilidade, X12 - Rácio de Cobertura de Juros, X13 - Rotação do Ativo, X14 - Rácio Capital Próprio-Vendas, X15 - Dimensão e X16 - Idade.

Comparando o Gráfico 1 com os coeficientes *standardized* representados na Tabela 14, conclui-se que as variáveis Rácio de Fundo de Maneio, Estrutura de Endividamento, Rotação do Ativo, Rácio Passivo-Capital Próprio, Rentabilidade Económica, Idade, Rentabilidade do Ativo, Rentabilidade das Vendas e Margem EBIT são as que apresentam maior peso e que mais se

diferenciam entre grupos, sendo que a zona de *cutoff* convencionada é de cerca de 30% (ver, p.e., SAGE, 2015, p. 600). As variáveis Rentabilidade do Ativo e Rentabilidade das Vendas, Liquidez Geral e Rácio Capital Próprio-Vendas, como se verifica sobrepõem-se, o que dificulta a diferenciação entre grupos.

Além da função discriminante para a amostra total, é também possível estimar o peso de cada variável na função discriminante consoante o grupo (saudável/insolvente), denominada por *estat classification*. Tal como se observa na Tabela 15 as variáveis com maior “peso” são a Estrutura de Endividamento, a Rentabilidade Económica e do Ativo, a Rentabilidade das Vendas, Margem EBIT e o Rácio de Endividamento.

Tabela 15 - Função Discriminante por grupo

Variável	Status = 0 Empresas saudáveis	Status = 1 Empresas insolventes
Liquidez Geral	1,786	1,800
Liquidez Imediata	3,496	3,461
Rácio de Fundo de Maneio	5,581	5,281
Estrutura de Endividamento	25,913	26,312
Rentabilidade Económica	117,317	118,654
Rentabilidade do Ativo	-107,477	-109,560
Margem EBIT	-27,779	-23,551
Rentabilidade das Vendas	21,077	18,436
Rácio Passivo – Capital Próprio	0,111	0,129
Rácio de Endividamento	28,505	28,527
Solvabilidade	-0,594	-0,547
Rácio de Cobertura de Juros	0,000	0,000
Rotação do Ativo	3,993	4,091
Rácio Capital Próprio – Vendas	7,670	7,603
Dimensão	8,814	8,878
Idade	3,355	3,568
Constante	-60,426	-62,004

A Tabela 15 apresenta a estimação do peso de cada variável na função discriminante por grupo de empresas (saudável/insolvente) para todas as variáveis explicativas. As variáveis são as descritas no Subcapítulo 5.2.2.

O próximo passo da ADM é a análise de correlação de Pearson, entre grupos em função das variáveis explicativas, designada por *Structure Matrix*. Os coeficientes da Tabela 16 representam o grau de correlação entre o peso de cada variável e a função discriminante.

Tabela 16 - *Structure Matrix*

Variável	Coefficientes
Liquidez Geral	0,006
Liquidez Imediata	-0,039
Rácio de Fundo de Maneio	0,152
Estrutura de Endividamento	-0,212
Rentabilidade Económica	-0,215
Rentabilidade do Ativo	-0,187
Margem EBIT	-0,390
Rentabilidade das Vendas	-0,338
Rácio Passivo – Capital Próprio	-0,407
Rácio de Endividamento	0,096
Solvabilidade	-0,178
Rácio de Cobertura de Juros	0,180
Rotação do Ativo	-0,131
Rácio Capital Próprio – Vendas	-0,054
Dimensão	-0,254
Idade	-0,479

A Tabela 16 mostra os coeficientes que medem a correlação entre cada variável explicativa e a função discriminante. As variáveis são as descritas no Subcapítulo 5.2.2.

Neste passo da ADM pretende-se identificar quais os coeficientes que apresentam o maior valor em termos absolutos, sendo 30% considerada a zona de *cutoff*. As variáveis explicativas que cumprem este critério são: Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio, Idade e Dimensão (embora esta última apresente um valor ligeiramente inferior à zona de *cutoff*). Os resultados desta análise reforçam os resultados estimados na análise anterior (Tabela 14), à exceção das variáveis Rácio de Fundo de Maneio, Estrutura de Endividamento, Rentabilidade Económica e Rotação do Ativo que têm um menor coeficiente de correlação com a função discriminante.

Por último, a análise ANOVA – *Analysis of Variance* - conclui as técnicas canónicas da metodologia ADM. Esta análise tem como objetivo a avaliação da seleção das variáveis com base no teste estatístico F, tendo como hipótese nula a inexistência de diferenças significativas relativamente ao poder discriminativo das variáveis entre grupos, sendo a hipótese nula rejeitada com um grau de significância ao nível de 5%. A análise ANOVA é a técnica estatística que permite

° Ver, p.e., SAGE (2015, p. 600).

avaliar percepções sobre as médias de populações. A análise visa fundamentalmente verificar se existe uma diferença significativa entre as médias dos grupos de tratamento e de controlo e se as variáveis explicativas exercem influência na variável dependente. A Tabela 17 apresenta os resultados.

Tabela 17 - Análise ANOVA

Variáveis	Teste F	p-value
Liquidez Geral	0,002	0,969
Liquidez Imediata	0,068	0,795
Rácio de Fundo de Maneio	1,041	0,308
Estrutura de Endividamento	2,028	0,154
Rentabilidade Económica	2,079	0,149
Rentabilidade do Ativo	1,585	0,208
Margem EBIT	6,843	0,009
Rentabilidade das Vendas	5,139	0,023
Rácio Passivo – Capital Próprio	7,455	0,006
Rácio de Endividamento	0,419	0,518
Solvabilidade	1,425	0,233
Rácio de Cobertura de Juros	1,469	0,226
Rotação do Ativo	0,776	0,378
Rácio Capital Próprio – Vendas	0,131	0,718
Dimensão	2,901	0,089
Idade	10,351	0,001

A Tabela 17 apresenta os resultados da análise ANOVA para todas as variáveis explicativas. As variáveis são as descritas no Subcapítulo 5.2.2.

Com base na evidência apresentada na Tabela 17, à exceção das variáveis Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio e Idade, todas as outras não apresentam capacidade explicativa da situação de insolvência.

Em suma, conclui-se que comparando a classificação observada e a estimada com base nas variáveis explicativas usadas neste estudo, 57,00% das empresas classificadas como saudáveis estão corretamente classificadas e 43,00% classificadas como saudáveis deveriam ser consideradas insolventes. Os resultados da análise *standardized* mostraram que as variáveis Rácio de Fundo de Maneio, Estrutura de Endividamento, Rotação do Ativo, Rácio Passivo-Capital Próprio, Rentabilidade Económica, Idade, Rentabilidade do Ativo, Rentabilidade das Vendas e Margem EBIT são os preditores mais eficientes da situação de insolvência. Os resultados da correlação de

Pearson identificaram as variáveis Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio, Dimensão e Idade com maior correlação na função discriminante. Por fim, os resultados da análise ANOVA sugerem que as variáveis Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio e Idade não devem ser excluídas, uma vez que o *p-value* associado ao teste F é inferior a 5%.

Por fim, os resultados apontam que as variáveis que mais se destacaram nas diferentes etapas da ADM devido ao seu poder explicativo foram as seguintes: Margem EBIT, Rentabilidade das Vendas, Rácio Passivo-Capital Próprio, Dimensão e Idade.

6.2. Análise Logística

A ADM foi colocada em causa por vários autores, sendo Ohlson (1980) um dos maiores críticos desta metodologia, que propôs um modelo de previsão de falência com base na análise logística, apresentando resultados mais robustos do que os anteriores. A ADM tem como desvantagem o facto de os seus pressupostos serem bastante restritivos. Já a análise *Logit* relaxa parte dos pressupostos da ADM, não exigindo que as variáveis independentes apresentem uma distribuição normal e que os grupos tenham matrizes de variância-covariância iguais, ou que os grupos sejam discretos e identificáveis. É ainda de salientar que a ADM é, apenas, uma técnica de classificação, enquanto o modelo *Logit* analisa uma relação causal.

Seguindo a intuição de Ohlson (1980), o objetivo da análise logística, neste estudo, é apresentar um modelo probabilístico com capacidade de prever a situação de insolvência até 3 anos antes desse evento. Assim, neste Subcapítulo serão percorridas todas as etapas necessárias para atingir esse objetivo.

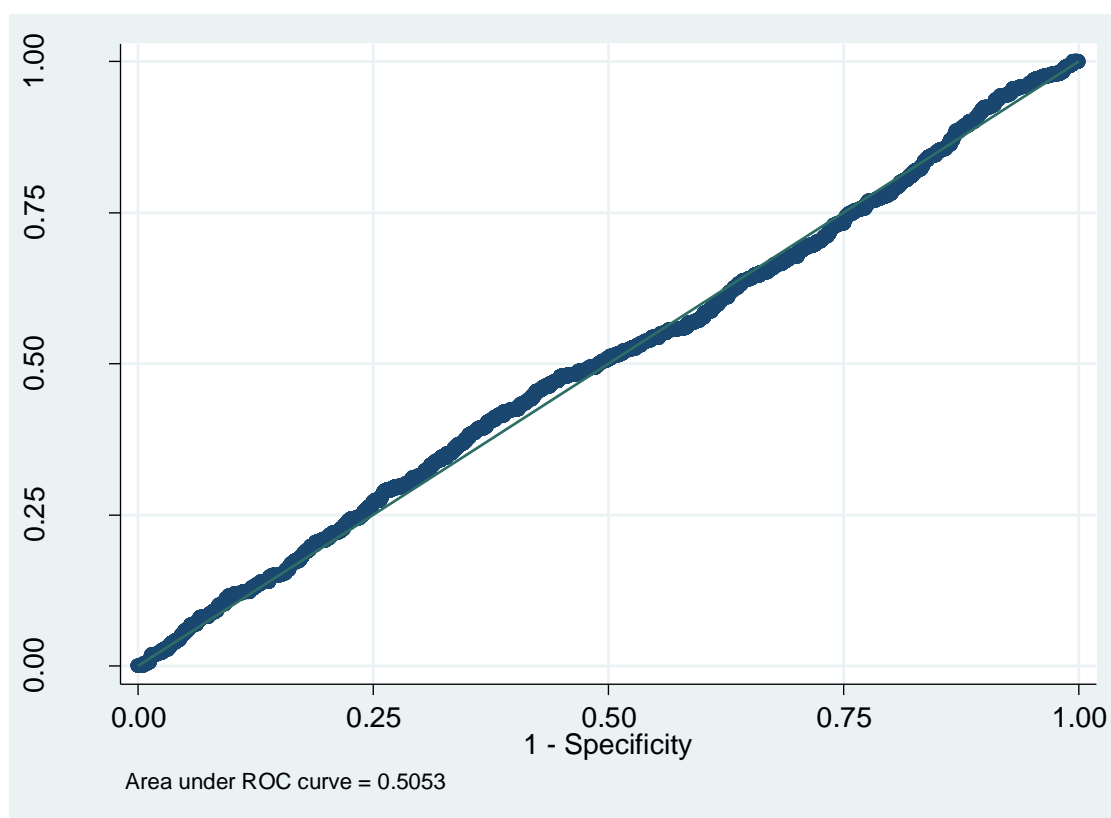
O primeiro passo da análise logística é semelhante ao da ADM, ou seja: a classificação das empresas com base no seu status (insolvente/saudável). Esta representação é efetuada através de um gráfico que ilustra o desempenho de um sistema classificador binário à medida que o seu limiar de discriminação varia de acordo com as variáveis explicativas incluídas no modelo, também designadas por *covariates*. Por outras palavras, esta análise permite estudar a variação da sensibilidade e especificidade, para diferentes níveis de *cutoff*, que diferencia o status entre insolvente/saudável.

O gráfico 2, denominado de ROC (*Receiver Operating Characteristic Curve*), corresponde a essa representação. A curva ROC apresenta a taxa de casos verdadeiros positivos (*Sensitivity*) versus a taxa de casos falsos positivos em diferentes cenários, sendo que a taxa de verdadeiros

negativos é, também, designada por *Specificity*, conforme demonstrado na Tabela 13. Em termos gráficos, a curva ROC é um gráfico de pares “xx” e “yy” que correspondem a $(1 - Specificity)$ e à *Sensitivity*, respetivamente, num plano designado por plano ROC unitário (sendo que cada uma dessas medidas representam uma probabilidade de acontecimento, e, por conseguinte, variam entre zero e um). Assim, de acordo com estas duas medidas, o objetivo é encontrar o ponto de *cutoff* que minimize o risco de se incorrer em erros.

A “*area under ROC curve*” significa a precisão com que a variável *Status* identifica de forma verdadeira as empresas insolventes. Neste caso, obteve-se uma precisão de 50,53%.

Gráfico 2 - ROC (*Receiver Operating Characteristic Curve*)



O Gráfico 2 representa a classificação das empresas em função do seu *Status* – insolvente/saudável – com base num modelo *Logit*, onde estão incluídas as variáveis explicativas, antes do processo de estandardização das variáveis.

A análise ROC sucede a estimação do modelo *Logit*, definido na equação (1).

$$Prob(Insolvente = 1) = \beta'Z_{i,t-1} + \lambda_j + \gamma_t + \varepsilon_{it} \quad (1)$$

Onde a variável dependente é o *status* da empresa, sendo igual a 1 quando a empresa se

encontra em situação de insolvência e 0 caso contrário. O vetor $Z_{i,t-1}$ inclui as variáveis explicativas desfasadas 1 período. Os parâmetros λ_j, γ_t são incluídos para controlarem a heterogeneidade não observada (p.e., neste caso concreto, relativa à indústria e/ou localização) e à variação temporal, respetivamente. Já o termo ε_{it} representa o erro da regressão.

Após a identificação das observações de cada grupo, com base nas suas características definidas pelas variáveis explicativas, conforme observado no Gráfico 2, a próximo passo é a estandardização dos coeficientes da equação (1). Este processo permite melhorar a comparação entre a magnitude dos coeficientes da estimação do modelo *Logit* e o seu poder explicativo.

A Tabela 18 apresenta os resultados da estimação da equação (1) antes do processo de estandardização dos coeficientes.

Tabela 18 – Resultados antes do Processo de Estandarização das Variáveis

Variável dependente: Insolvente	Coefficientes	Desvio Padrão	Z-test	p-value
Liquidez Geral	-0,065	0,049	-1,330	0,182
Liquidez Imediata	-0,189	0,093	-2,030	0,043
Rácio de Fundo de Maneio	-0,496	0,233	-2,130	0,033
Estrutura de Endividamento	-0,691	0,213	-3,240	0,001
Rentabilidade Económica	-3,478	2,984	-1,170	0,244
Rentabilidade do Ativo	1,606	3,205	0,500	0,616
Margem EBIT	4,947	1,857	2,660	0,008
Rentabilidade das Vendas	-2,039	1,896	-1,080	0,282
Rácio Passivo – Capital Próprio	0,014	0,005	2,740	0,006
Rácio de Endividamento	-1,773	0,277	-6,400	0,000
Solvabilidade	0,003	0,077	0,040	0,968
Rácio de Cobertura de Juros	-4,730	3,690	-1,280	0,200
Rotação do Ativo	-0,097	0,056	-1,730	0,083
Rácio Capital Próprio – Vendas	-0,587	0,194	-3,020	0,002
Dimensão	-0,290	0,039	-7,390	0,000
Idade	0,053	0,058	0,910	0,362

Observações = 32 208

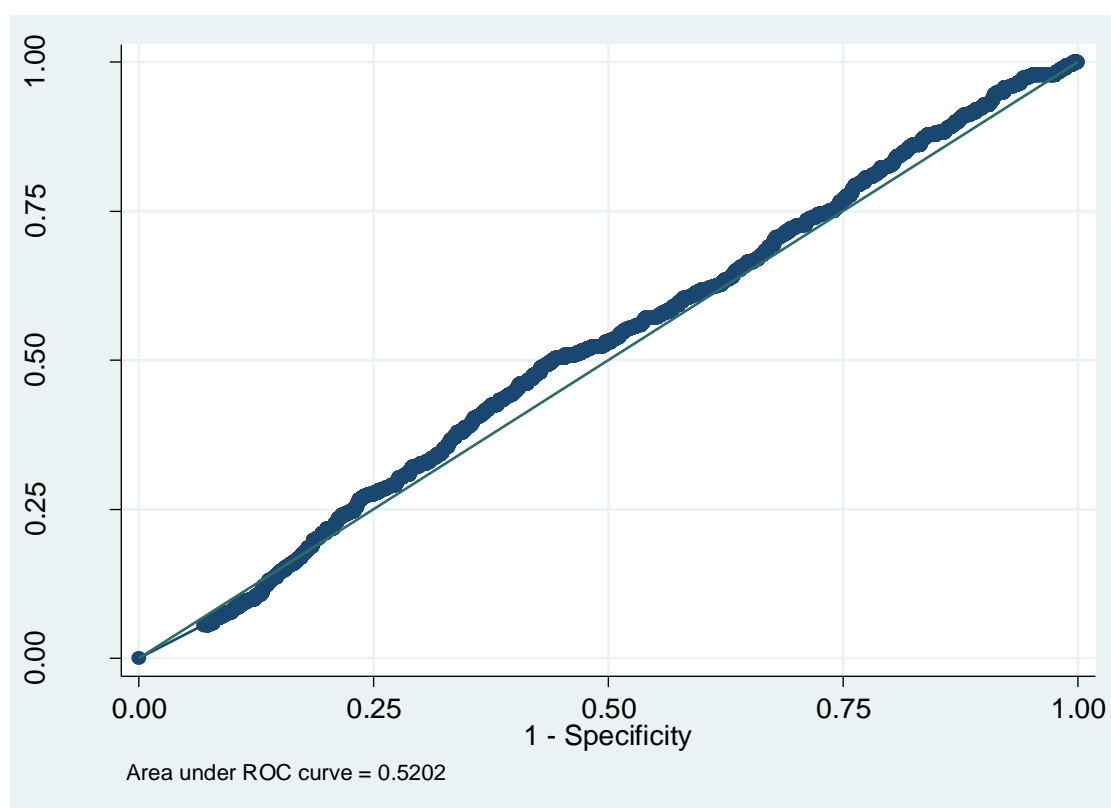
A Tabela 18 apresenta os coeficientes da estimação da equação (1), antes do processo de estandardização. A variável dependente assume caráter binário identificando as empresas com *Status* 1 se insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito.

Os coeficientes reportados na Tabela 18 são apresentados como *log-odds*. Usando a

variável Liquidez Imediata como um exemplo, o seu coeficiente é de -0,189, o que significa que quando a Liquidez Imediata diminui 1%, a empresa aumenta em 45,3% a probabilidade de insolvência, sendo que a $P(y=1)=\exp(-0.189)/(1 + \exp(-0.189))$, mantendo tudo o resto constante⁷.

Após o processo de estandardização das variáveis, observa-se um incremento na curva ROC. Neste caso, obteve-se uma precisão de 52,02%.

Gráfico 3 - ROC (*Receiver Operating Characteristic Curve*)



O Gráfico 3 representa a classificação das empresas em função do seu *Status* – insolvente/saudável – com base num modelo *Logit*, onde estão incluídas as variáveis explicativas, após o processo de estandardização das variáveis.

A Tabela 19 apresenta os resultados da estimação da equação (1) após o processo de estandardização.

Comparando as Tabelas 18 e 19 antes e depois do processo de estandardização, constata-se que as diferenças na magnitude e na significância dos coeficientes são notórias. Como se pode observar, após a estandardização das variáveis, todas elas são estatisticamente significativas ao nível de 1%, o que mostra que têm um elevado poder explicativo.

⁷ Uma forma alternativa de interpretação do coeficiente é que para um aumento de uma unidade na variável Liquidez Imediata, espera-se uma diminuição de 0,189 na probabilidade logarítmica da variável dependente Insolvente, mantendo todas as outras variáveis independentes constantes.

Tabela 19 – Resultados após o Processo de Estandarização das Variáveis

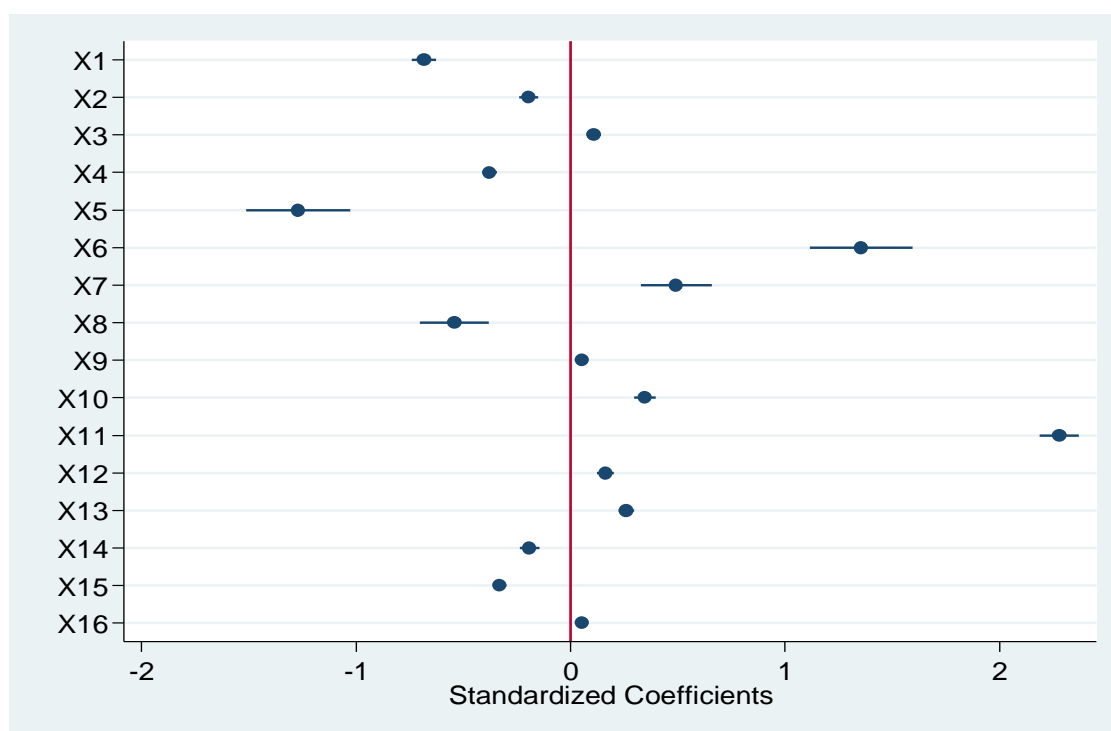
Variável dependente: Insolvente	Coefficientes	Desvio Padrão	Z-test	p-value
Liquidez Geral	-0,685	0,029	-23,670	0,000
Liquidez Imediata	-0,195	0,022	-8,920	0,000
Rácio de Fundo de Maneio	0,108	0,017	6,250	0,000
Estrutura de Endividamento	-0,379	0,017	-21,760	0,000
Rentabilidade Económica	-1,270	0,123	-10,310	0,000
Rentabilidade do Ativo	1,354	0,122	11,090	0,000
Margem EBIT	0,493	0,084	5,830	0,000
Rentabilidade das Vendas	-0,541	0,082	-6,630	0,000
Rácio Passivo – Capital Próprio	0,053	0,013	4,220	0,000
Rácio de Endividamento	0,346	0,026	13,410	0,000
Solvabilidade	2,278	0,046	48,990	0,000
Rácio de Cobertura de Juros	0,163	0,021	7,930	0,000
Rotação do Ativo	0,260	0,019	13,720	0,000
Rácio Capital Próprio – Vendas	-0,191	0,024	-8,090	0,000
Dimensão	-0,330	0,016	-20,840	0,000
Idade	0,054	0,014	3,790	0,000

Observações = 32 208

A Tabela 19 apresenta os coeficientes da estimação da equação (1), após o processo de estandarização. A variável dependente assume caráter binário identificando as empresas com *Status* 1 se insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito.

Tal como acontece com a estimação dos coeficientes *standardized* na análise discriminante, este processo permite mitigar as diferenças de escala entre as variáveis, o que conduz a resultados mais robustos. Comparando os resultados das Tabelas 18 e 19, os coeficientes das variáveis Liquidez Geral, Rentabilidade Económica, Rentabilidade do Ativo, Rentabilidades das Vendas, Solvabilidade, Rácio de Cobertura de Juros e Idade ganharam significância estatística. O gráfico seguinte apresenta a significância dos coeficientes, que tem como objetivo identificar quais os coeficientes que se encontram mais afastados de zero, ou seja, que são estatisticamente diferentes de zero.

Gráfico 4 - Significância dos Coeficientes Estandarizados



O Gráfico 4 representa a significância estatística das variáveis explicativas, após o processo de estandardização. As variáveis são as descritas no Subcapítulo 5.2.2. e são as seguintes: X1 - Liquidez Geral, X2 - Liquidez Imediata, X3 - Rácio de Fundo de Maneio, X4 - Estrutura de Endividamento, X5 - Rentabilidade Económica, X6 - Rentabilidade do Ativo, X7 - Margem EBIT, X8 - Rentabilidade das Vendas, X9 - Rácio Passivo-Capital Próprio, X10 - Rácio de Endividamento, X11 - Solvabilidade, X12 - Rácio de Cobertura de Juros, X13 - Rotação do Ativo, X14 - Rácio Capital Próprio-Vendas, X15 - Dimensão e X16 - Idade.

A observação do Gráfico 4 conduz à conclusão de que praticamente todos os coeficientes se encontram afastados de zero, à exceção das variáveis Rácio de Fundo de Maneio, Rácio Passivo-Capital Próprio e Idade.

Tal como relatado anteriormente, as críticas de Ohlson (1980) no que concerne à validade e robustez da análise discriminante levou ao objetivo de apresentar um modelo de previsão de insolvência suportado num modelo logístico. Neste contexto, após o processo de estandardização, todas as variáveis apresentam significância estatística, ou por outras palavras, aparentam ter capacidade preditiva da situação de insolvência, dado que a inferência estatística aponta para que sejam estatisticamente diferentes de zero.

Após este processo, as variáveis selecionadas para integrarem o modelo final, foram as seguintes: Estrutura de Endividamento, Margem EBIT, Rácio Passivo – Capital Próprio, Rácio de Cobertura de Juros, Rotação do Ativo, Dimensão e Idade, uma vez que foram as que apresentaram melhores resultados, ao longo das análises anteriores, quando analisadas em simultâneo.

6.3. Propensity Score Matching

Como já referido nos Subcapítulos anteriores, a análise *Logit* apresenta vantagens em relação à ADM, no entanto, a análise *Logit* relaxa alguns pressupostos que poderão ter impacto nos resultados. Uma dessas limitações é o facto de os grupos de empresas insolventes/saudáveis serem demasiado desproporcionais; outra limitação neste tipo de estudos é o facto de os grupos de empresas apresentarem características significativamente diferentes, como é possível observar na Tabela 11 no Subcapítulo 5.4. Uma solução para ultrapassar estas limitações, ou seja, a desproporcionalidade de observações entre empresas insolventes e saudáveis e a suas diferenças estruturais, é adotar técnicas de *matching*. Desta forma, neste estudo será aplicada a técnica de *Propensity Score Matching*.

Em conformidade com a descrição efetuado no Subcapítulo 4.2, a primeira etapa do *Propensity Score Matching* consiste na estimação do *propensity score*, ou seja, a probabilidade de uma empresa ser designada para o grupo de tratamento, que neste caso significa ser insolvente, através de um modelo *Probit* estimado a partir de um conjunto de variáveis explicativas da situação de insolvência. Esta análise foi realizada assumindo como *covariates* as variáveis explicativas usadas nos Subcapítulos anteriores. A Tabela 20 apresenta os resultados da estimação do *propensity score*.

Relativamente Tabela 20 que diz respeito à estimação do *propensity score*, observa-se que os coeficientes das variáveis Margem EBIT, Rácio Passivo-Capital Próprio, Rotação do Ativo e Idade são estatisticamente significativos ao nível de pelo menos 10%.

Um dos passos implícitos no processo de *Propensity Score Matching* é dividir a amostra em blocos, que é semelhante a dividir a amostra em subgrupos com características comuns para se formarem pares (*matches*) entre as observações de empresas insolventes e saudáveis.

Tabela 20 – Estimação do *Propensity Score*

Variável dependente: Insolvente	Coefficientes	Desvio Padrão	Z-test	p-value
Liquidez Geral	0,003	0,019	0,170	0,868
Liquidez Imediata	-0,017	0,037	-0,450	0,649
Rácio de Fundo de Maneio	-0,109	0,091	-1,200	0,232
Estrutura de Endividamento	0,155	0,095	1,640	0,102
Rentabilidade Económica	0,582	1,194	0,490	0,626
Rentabilidade do Ativo	-0,962	1,290	-0,750	0,456
Margem EBIT	1,529	0,843	1,810	0,070
Rentabilidade das Vendas	-0,728	0,867	-0,840	0,401
Rácio Passivo – Capital Próprio	0,007	0,002	3,390	0,001
Rácio de Endividamento	-0,012	0,093	-0,130	0,896
Solvabilidade	0,020	0,026	0,740	0,458
Rácio de Cobertura de Juros	0,000	0,000	-1,470	0,141
Rotação do Ativo	0,035	0,020	1,760	0,078
Rácio Capital Próprio – Vendas	-0,044	0,066	-0,660	0,507
Dimensão	0,025	0,019	1,320	0,187
Idade	0,085	0,024	3,570	0,000
Constante	-2,695	0,187	-14,390	0,000

Observações = 32 208

A Tabela 20 apresenta as estimativas dos coeficientes do modelo *Probit* usado para estimar o *Propensity Score*. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito.

No segundo passo do processo de *matching*, aplicamos a técnica de *matching* com *nearest neighbor and caliper*, que corresponde a um intervalo, onde a probabilidade de *propensity score* pode variar. Aplicar o *caliper matching* significa que uma empresa do grupo de controlo será combinada com uma empresa de tratamento que esteja dentro desse intervalo. O *caliper* adequado foi estimado de acordo com Cochran and Rubin, (1973) e corresponde a 25% do desvio padrão do *propensity score*.

A Tabela 21 apresenta os resultados da qualidade final do processo de *matching*.

Tabela 21 – Resultados do teste de qualidade do *matching*

Variável	Grupo	Média		t-test	p-value
		Tratamento	Controlo		
Liquidez Geral	<i>Unmatched</i>	2,207	2,210	-0,040	0,969
	<i>Matched</i>	2,161	2,170	-0,090	0,931
Liquidez Imediata	<i>Unmatched</i>	0,430	0,422	0,260	0,795
	<i>Matched</i>	0,421	0,434	-0,310	0,758
Rácio de Fundo de Maneio	<i>Unmatched</i>	0,236	0,246	-1,020	0,308
	<i>Matched</i>	0,240	0,239	0,110	0,915
Estrutura de Endividamento	<i>Unmatched</i>	0,610	0,595	1,420	0,154
	<i>Matched</i>	0,607	0,614	-0,470	0,635
Rentabilidade Económica	<i>Unmatched</i>	0,050	0,044	1,440	0,149
	<i>Matched</i>	0,051	0,045	1,000	0,318
Rentabilidade do Ativo	<i>Unmatched</i>	0,026	0,021	1,260	0,208
	<i>Matched</i>	0,027	0,022	0,840	0,400
Margem EBIT	<i>Unmatched</i>	0,044	0,034	2,620	0,009
	<i>Matched</i>	0,043	0,039	0,690	0,488
Rentabilidade das Vendas	<i>Unmatched</i>	0,021	0,013	2,270	0,023
	<i>Matched</i>	0,020	0,019	0,260	0,795
Rácio Passivo – Capital Próprio	<i>Unmatched</i>	3,995	3,103	2,730	0,006
	<i>Matched</i>	3,136	2,741	1,110	0,265
Rácio de Endividamento	<i>Unmatched</i>	0,676	0,684	-0,650	0,518
	<i>Matched</i>	0,672	0,671	0,070	0,944
Solvabilidade	<i>Unmatched</i>	0,851	0,794	1,190	0,233
	<i>Matched</i>	0,818	0,840	-0,300	0,761
Rácio de Cobertura de Juros	<i>Unmatched</i>	947,210	1866,600	-1,210	0,226
	<i>Matched</i>	995,750	1706,100	-0,910	0,364
Rotação do Ativo	<i>Unmatched</i>	1,484	1,445	0,880	0,378
	<i>Matched</i>	1,422	1,568	-2,350	0,019
Rácio Capital Próprio – Vendas	<i>Unmatched</i>	0,366	0,360	0,360	0,718
	<i>Matched</i>	0,358	0,345	0,530	0,594
Dimensão	<i>Unmatched</i>	6,824	6,744	1,700	0,089
	<i>Matched</i>	6,814	6,756	0,860	0,389
Idade	<i>Unmatched</i>	2,810	2,700	3,220	0,001
	<i>Matched</i>	2,776	2,797	-0,410	0,681

A Tabela 21 apresenta as diferenças entre grupos de empresas de tratamento e de controlo que foram *Matched*, e os grupos de empresas de tratamento e de controlo para as quais não foi possível encontrar um *match*, pertencendo ao grupo *Unmatched*. O nível de significância a hipótese nula do teste é de 5%. As variáveis explicativas são as descritas no Subcapítulo 5.2.2.

Os resultados da Tabela 21 revelam que tanto para os grupos de tratamento e de controlo *Unmatched* como *Matched*, não existem diferenças significativas, à exceção das variáveis Rácio Passivo-Capital Próprio e Rácio de Cobertura de Juros, o que indica que as médias das variáveis entre os grupos de tratamento e de controlo são semelhantes após o *matching*, à exceção das variáveis enumeradas anteriormente.

A qualidade do *matching* está assegurada quando não se observam diferenças estatisticamente significativas, entre os grupos de tratamento (empresas insolventes) e de controlo (empresas saudáveis) no grupo *Matched*. Como se pode verificar na Tabela 21, o *p-value* é sempre superior ao nível de significância de 5%, logo não existem diferenças estatisticamente significativas, entre os grupos.

Dada a evidência da qualidade do processo de *matching*, a análise da estimação do *propensity score* será repetida para um grupo restrito de variáveis, que apresentaram maior preditivo ao longo das análises anteriores, entre as quais: Estrutura de Endividamento, Margem EBIT, Rácio Passivo – Capital Próprio, Rácio de Cobertura de Juros, Rotação do Ativo, Dimensão e Idade.

A Tabela 22 dá-nos os resultados para uma nova estimação do *propensity score* considerando um modelo reduzido às variáveis mais relevantes.

Tabela 22 – Estimação do *Propensity Score* para um modelo restrito

Variável dependente: Insolvente	Coefficientes	Desvio Padrão	Z	p-value
Estrutura de Endividamento	0,148	0,076	1,960	0,050
Margem EBIT	0,603	0,210	2,870	0,004
Rácio Passivo – Capital Próprio	0,007	0,002	3,380	0,001
Rácio de Cobertura de Juros	0,000	0,000	-1,470	0,142
Rotação do Ativo	0,038	0,018	2,130	0,033
Dimensão	0,026	0,018	1,430	0,154
Idade	0,081	0,023	3,470	0,001
Constante	-2,695	0,151	-17,820	0,000

Observações = 32 298

A Tabela 22 apresenta as estimativas dos coeficientes do modelo *Probit* usado para estimar o *Propensity Score* incluindo apenas as variáveis com maior capacidade preditiva. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito.

Os resultados da Tabela 22 demonstram que os coeficientes das variáveis são, em média,

estatisticamente significativos ao nível de pelo menos de 10%, à exceção das variáveis Rácio de Cobertura de Juros e Dimensão.

À semelhança da Tabela 21, a Tabela 23 apresenta os resultados da qualidade do *matching*.

Os resultados obtidos na Tabela 23 são similares aos observados na Tabela 21, ou seja, os grupos de tratamento e de controlo *Unmatched* e *Matched* não apresentam diferenças significativas, o que indica que as médias das variáveis entre os grupos de tratamento e de controlo são semelhantes, após o *matching*.

Tabela 23 – Resultados do teste de qualidade do *matching*

Variável	Grupo	Média		t-test	p-value
		Tratamento	Controlo		
Estrutura de Endividamento	<i>Unmatched</i>	0,610	0,595	1,490	0,135
	<i>Matched</i>	0,605	0,632	-1,850	0,064
Margem EBIT	<i>Unmatched</i>	0,044	0,034	2,630	0,009
	<i>Matched</i>	0,043	0,040	0,790	0,430
Rácio Passivo – Capital Próprio	<i>Unmatched</i>	3,974	3,105	2,650	0,008
	<i>Matched</i>	3,085	2,516	1,700	0,090
Rácio de Cobertura de Juros	<i>Unmatched</i>	945,550	1876,200	-1,220	0,221
	<i>Matched</i>	1013,200	1645,800	-0,740	0,459
Rotação do Ativo	<i>Unmatched</i>	1,485	1,444	0,910	0,363
	<i>Matched</i>	1,467	1,474	-0,120	0,903
Dimensão	<i>Unmatched</i>	6,825	6,742	1,780	0,075
	<i>Matched</i>	6,829	6,875	-0,680	0,498
Idade	<i>Unmatched</i>	2,808	2,699	3,200	0,001
	<i>Matched</i>	2,773	2,791	-0,330	0,739

A Tabela 23 apresenta as diferenças entre grupos de empresas de tratamento e de controlo que foram *Matched*, e os grupos de empresas de tratamento e de controlo para as quais não foi possível encontrar um *match*, pertencendo ao grupo *Unmatched*. As variáveis explicativas são as incluídas no modelo reduzido na Tabela 22. O nível de significância a hipótese nula do teste é de 5%. As variáveis explicativas são as descritas no Subcapítulo 5.2.2.

No entanto, a qualidade do modelo *Probit* que estima o *propensity score* é testada com recurso ao teste de máxima verossimilhança (*Likelihood Ratio* - LR) χ^2 , que testa a adequação (*goodness-of-fit*) desse modelo *Probit*. Se o *propensity score* for o mais adequado, o nível de significância conjunta dos coeficientes deve ser zero ou próximo de zero (para um nível de significância estatística de 5%). Tal como podemos observar na Tabela 24, esta condição é

verificada pela análise decorrente do valor do teste chi2 (*p-value* de 29,1%).

Tabela 24 – *Goodness-of-fit* do *propensity score*

Grupo	Pseudo R2	LR chi2	<i>p-value</i>
<i>Unmatched</i>	0,007	38,980	0,000
<i>Matched</i>	0,006	8,490	0,291

A Tabela 20 representa, de forma resumida, o rácio de máxima verosimilhança (LR), cujo teste estatístico associado é o chi2. A hipótese nula do teste é rejeitada com um nível de significância inferior a 5%.

Em síntese, o método *Propensity Score Matching* (PSM) foi estimado usando como *covariates* as variáveis selecionadas para integrarem o modelo final, que são as seguintes: Estrutura de Endividamento, Margem EBIT, Rácio Passivo – Capital Próprio, Rácio de Cobertura de Juros, Rotação do Ativo, Dimensão e Idade. Os testes revelaram que a qualidade do *matching* está assegurada, pelo que os resultados suportam a hipótese de que o *propensity score* estimado reflete a probabilidade dessas variáveis explicativas corresponderem às *covariates* adequadas para integrarem o modelo final de previsão de insolvência.

6.4. Modelo Final

6.4.1. Resultados do Modelo *Logit* para dois e três anos antes da Insolvência

Com base nos resultados das análises anteriores, nesta fase o objetivo é definir o modelo empírico de previsão de insolvência que inclui o conjunto de variáveis que foram selecionadas nessas secções. Nesta secção, optou-se por apenas apresentar resultados para dois e três anos antes, uma vez que, os resultados para um ano antes da insolvência diferem pouco dos resultados para dois anos antes do evento⁸. Assim, na Tabela 25 são apresentadas as estimativas dos coeficientes do modelo *Logit* definido na equação (2), cuja variável dependente é a *dummy* que identifica as empresas insolventes e as variáveis explicativas são as seguintes: Estrutura de Endividamento, Margem EBIT, Rácio Passivo – Capital Próprio, Rácio de Cobertura de Juros, Rotação do Ativo, Dimensão e Idade, como determinantes na avaliação da insolvência empresarial.

⁸ Como não existem diferenças consideráveis dos dados para um e dois anos antes da insolvência, optou-se por apresentar apenas os resultados para dois e três anos antes do evento.

$$\begin{aligned}
\text{Prob}(\text{Insolvente} = 1) = & \alpha + \beta_1 \text{Estrutura de Endividamento} + \\
& \beta_2 \text{Margem EBIT} + \beta_3 \text{Rácio Passivo} - \text{Capital Próprio} + \\
& \beta_4 \text{Rácio de Cobertura de Juros} + \beta_5 \text{Rotação do Ativo} + \beta_6 \text{Dimensão} + \\
& \beta_7 \text{Idade} + \lambda_j + \gamma_t + \varepsilon_{it}
\end{aligned}
\tag{2}$$

A Tabela 25 apresentada os resultados.

Tabela 25 – Resultados da estimação do modelo *Logit*

Variável dependente: Insolvente	2 anos antes	2 anos antes	3 anos antes	3 anos antes
	(coeficientes)	(<i>p-value</i>) z	(coeficientes)	(<i>p-value</i>) z
	(1)		(2)	
Estrutura de Endividamento	-0,779	(0,733)	-1,690	(0,521)
Margem EBIT	-3,720	(0,235)	-4,482	(0,147)
Rácio Passivo – Capital Próprio	0,011	(0,111)	0,008	(0,260)
Rácio de Cobertura de Juros	-0,000	(0,270)	-0,000	(0,374)
Rotação do Ativo	-0,639	(0,000)	-0,625	(0,001)
Dimensão	-1,194	(0,000)	-1,047	(0,000)
Idade	-0,501	(0,256)	-0,305	(0,573)
Constante	1,307	(0,611)	0,335	(0,896)
Significância Global				
Coeficientes: <i>p-value</i> (chi2-test)		(0,000)		(0,000)
Pseudo R2		0,113		0,098
Observações corretamente classificadas		82,58%		90,41%
Observações		31 737		31 736

A Tabela 25 reporta os resultados da estimação da equação (2), nomeadamente os coeficientes e o *p-value* associado a cada coeficiente. A variável dependente é binária, assumindo o valor de 1 quando a empresa é insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. É, ainda, apresentada a significância estatística conjunta para todas as variáveis (teste chi2), o Pseudo R2 e a percentagem de acerto do modelo, para dois e três anos antes da insolvência. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito. Os erros-padrão são *White-robust*.

Os resultados apresentados na Tabela 25 demonstram evidência que as variáveis explicativas incluídas no modelo têm no seu conjunto capacidade preditiva da situação de insolvência (conforme reportado pelo teste de significância conjunta chi2) a um nível de significância de 1% até 3 anos antes do evento. O conjunto das variáveis selecionadas é estatisticamente significativo e tem capacidade para explicar a variável dependente. Nos dois modelos (dois e três anos antes da insolvência), as estimativas das variáveis Rotação do Ativo e Dimensão são estatisticamente significativas a um nível de significância de 1%. De realçar, ainda,

que para além do contributo individual dos coeficientes para a explicação da situação de insolvência, é o contributo global do modelo que permite aferir a probabilidade de uma empresa entrar numa situação de insolvência.

Assim sendo, o modelo mostra ser bastante eficaz uma vez que, de modo geral, prevê mais de 82% das observações corretamente⁹, para dois anos antes do evento, chegando aos 90% para três anos antes.

Com o objetivo de testar a robustez do modelo, incluiu-se na equação (2) um conjunto de *dummies* para controlar para efeitos temporais, para efeitos de localização (NUTS2) e para efeitos de tipo de indústria (conforme as indústrias definidas no Anexo A do Apêndice). A Tabela 26 mostra os resultados.

Tabela 26 – Resultados da estimação do modelo *Logit* incluindo controlo para efeitos temporais, localização e indústria

Variável dependente: Insolvente	2 anos antes	2 anos antes	3 anos antes	3 anos antes
	(coeficientes)	(<i>p-value</i> z)	(coeficientes)	(<i>p-value</i> z)
	(1)		(2)	
Estrutura de Endividamento	-0,224	(0,922)	-0,983	(0,712)
Margem EBIT	-2,652	(0,493)	-3,982	(0,330)
Rácio Passivo – Capital Próprio	0,003	(0,816)	-0,002	(0,898)
Rácio de Cobertura de Juros	-0,000	(0,302)	-0,000	(0,638)
Rotação do Ativo	-0,448	(0,097)	-0,321	(0,432)
Dimensão	-0,965	(0,015)	-0,764	(0,024)
Idade	-0,680	(0,230)	-0,428	(0,592)
Constante	-0,022	(0,996)	-1,764	(0,704)
Significância Global				
Coeficientes: <i>p-value</i> (chi2-test)		(0,000)		(0,000)
Pseudo R2		0,228		0,244
Observações Corretamente Classificadas		91,52%		92,45%
Observações		9 532		7 507

A Tabela 26 reporta os resultados da estimação da equação (2), incluindo o controlo para efeitos temporais, localização e indústria, nomeadamente os coeficientes e o *p-value* associado a cada coeficiente. A variável dependente é binária, assumindo o valor de 1 quando a empresa é insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. É, ainda, apresentada a significância estatística conjunta para todas as variáveis (teste chi2), o Pseudo R2 e a percentagem de acerto do modelo, para dois e três anos antes da insolvência. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito. Os erros-padrão são *White-robust*.

⁹ De acordo com o estudo de Ohlson (1980), o número de observações corretamente classificadas é obtido definindo uma zona de *cutoff* do modelo, que corresponde ao valor mínimo observado dos erros tipo II (ver Tabela 13).

De um modo geral, controlando-se para efeitos temporais, de localização e de indústria não diminui a capacidade do modelo para explicar a situação de insolvência a um nível de significância de 1%; o teste chi2 para a significância conjunta das variáveis é estatisticamente significativo ao nível de 1%. Mais uma vez, a Tabela 26 apresenta resultados para dois e três anos antes do evento da insolvência. Desta forma, os resultados demonstram evidência que para dois anos antes da insolvência, as estimativas dos coeficientes das variáveis Rotação do Ativo e Dimensão são estatisticamente significativas para um nível de significância de 10 e 5%, respetivamente. Já para três anos antes da insolvência, a estimativa do coeficiente da variável Dimensão é estatisticamente significativa a um nível de significância de 5%. Em termos de eficácia no que respeita ao acerto, o modelo prevê mais de 90% das observações corretamente para 3 anos antes do evento.

6.4.2. Resultados do Modelo *Probit* para dois e três anos antes da Insolvência

De forma a ser possível comparar resultados utilizando diferentes metodologias, a equação (2) foi estimada com recurso a um modelo *Probit*. Nesta análise, foram usadas nas regressões as observações *matched* resultantes do processo de *Propensity Score Matching*.

Consistente com a análise anterior, os resultados obtidos na Tabela 27 demonstram que o conjunto das variáveis incluídas na equação (2) tem a capacidade preditiva para explicar a variável dependente a um nível de significância de 1%.

Tabela 27 – Resultados da estimação do modelo *Probit*

Variável dependente: Insolvente	2 anos antes	2 anos	3 anos antes	3 anos
	(coeficientes)	antes	(coeficientes)	antes
	(1)	(<i>p-value</i>) z	(2)	(<i>p-value</i>) z
Estrutura de Endividamento	-0,716	(0,764)	-1,661	(0,546)
Margem EBIT	-4,529	(0,127)	-5,001	(0,090)
Rácio Passivo – Capital Próprio	0,012	(0,189)	0,010	(0,404)
Rácio de Cobertura de Juros	-0,000	(0,396)	-0,000	(0,666)
Rotação do Ativo	-0,616	(0,000)	-0,594	(0,001)
Dimensão	-1,299	(0,000)	-1,134	(0,000)
Idade	-0,563	(0,161)	-0,369	(0,486)
Constante	3,961	(0,144)	2,914	(0,281)
Significância Global		(0,000)		(0,000)
Coeficientes: <i>p-value</i> (chi2-test)				
Pseudo R2	0,159		0,133	
Observações Corretamente Classificadas		81,72%		89,29%
Observações		4 502		4 501

A Tabela 27 apresenta os resultados da estimação da equação (2), recorrendo a um modelo *Probit*. São apresentados os coeficientes e o *p-value* associado a cada coeficiente. A variável dependente é binária, assumindo o valor de 1 quando a empresa é insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. É, ainda, apresentada a significância estatística conjunta para todas as variáveis (teste chi2), o Pseudo R2 e a percentagem de acerto do modelo, para dois e três anos antes da insolvência. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito.

À semelhança das Tabelas 25 e 26, a Tabela 27 apresenta os resultados para dois e três anos antes da insolvência. O modelo apresenta uma taxa de acerto superior a 80% para dois anos antes do evento e de, aproximadamente, 89% para três anos antes.

6.4.3. Modelo Final de Previsão de Insolvência

Com base nas variáveis selecionadas nas secções anteriores, as estimativas da regressão do modelo de previsão definido na equação (2) são apresentadas na Tabela 28, seguida da representação gráfica da respetiva curva ROC (*Receiver Operating Characteristic*), a qual relaciona as métricas de sensibilidade (casos que são verdadeiros positivos) com a especificidade (casos que são verdadeiros negativos) do modelo estimado.

De salientar mais uma vez que a probabilidade de insolvência é dada pelo somatório de todos os coeficientes no modelo, sendo que a significância conjunta de todos os coeficientes é de

1%. O modelo geral de previsão de insolvência apresenta uma percentagem de observações corretamente classificadas de 64,80%.

Tabela 28 – Resultados da estimação do modelo *Logit* final

Variável dependente: Insolvente	Coefficientes	Desvio Padrão	Z-test	p-value
Estrutura de Endividamento	0,355	0,190	1,870	0,062
Margem EBIT	1,562	0,586	2,670	0,008
Rácio Passivo – Capital Próprio	0,016	0,005	3,660	0,000
Rácio de Cobertura de Juros	-0,001	0,000	-1,470	0,143
Rotação do Ativo	0,100	0,043	2,300	0,022
Dimensão	0,065	0,048	1,370	0,171
Idade	0,218	0,075	2,930	0,003
Constante	-5,543	0,414	-13,390	0,000

Observações = 32 298

Significância Global (Prob > chi2): 0,000

Pseudo R2: 0,007

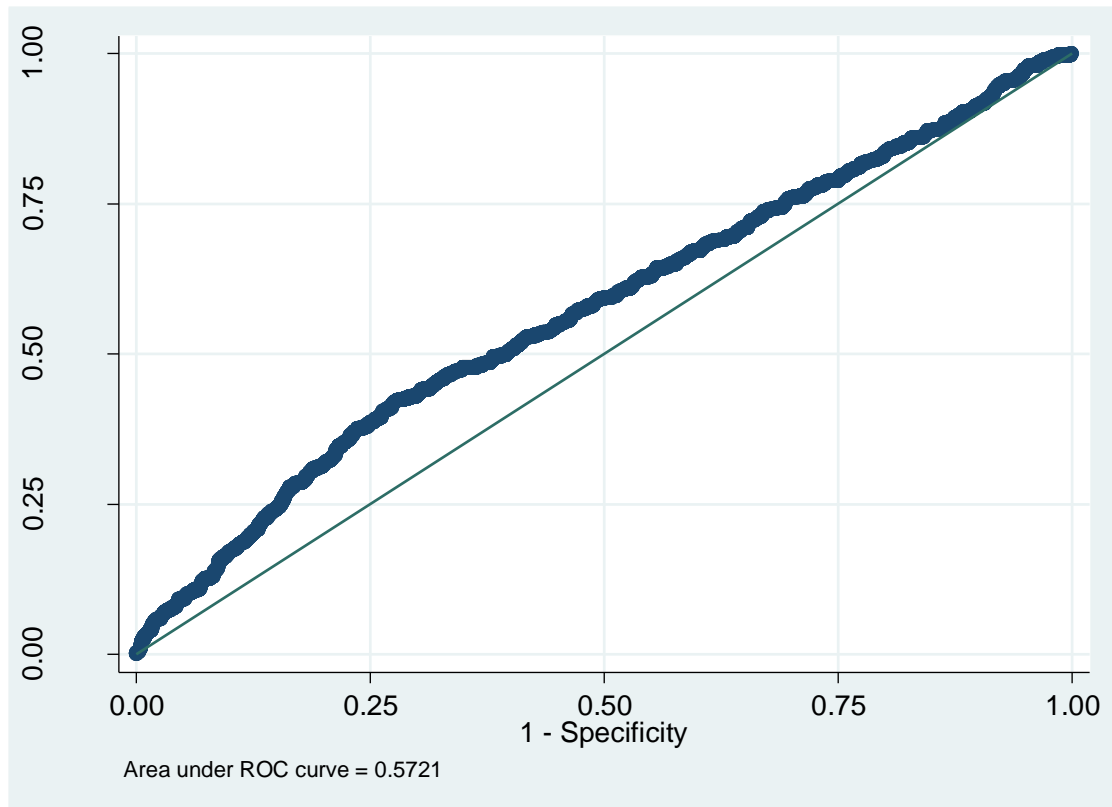
Observações Corretamente Classificadas: 64,80%

A Tabela 28 apresenta os coeficientes da estimação da equação (2). A variável dependente é uma *dummy* que identifica as empresas com status 1 se insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negrito. Os erros-padrão são *White-robust*.

O Gráfico 5 representa a curva ROC do modelo final; o modelo apresenta uma capacidade preditiva¹⁰ de 57,21%.

¹⁰ Embora possa ser considerado razoável, situa-se abaixo do valor estilizado de 70% (ver, p.e., Hosmer, Lemeshow & Sturdivant, 2013, p.177).

Gráfico 5 - Curva ROC (Modelo Final)



O Gráfico 5 apresenta a classificação das empresas em função do seu *Status* - insolvente/saudável - com base na equação (2), estimada com recurso a um modelo *Logit*.

O modelo final de previsão de insolvência é o seguinte:

$$\text{Score} = -5,543 + 0,355 \text{ Estrutura de Endividamento} + 1,562 \text{ Margem EBIT} + 0,016 \text{ Rácio Passivo} - \text{Capital Próprio} - 0,001 \text{ Rácio de Cobertura de Juro} + 0,100 \text{ Rotação do Ativo} + 0,065 \text{ Dimensão} + 0,218 \text{ Idade} \quad (3)$$

O modelo final de previsão de insolvência apresenta uma zona de *cutoff*¹ de 1,9%, ou seja, se o *score*¹ da equação (3) for superior a 0,019 a empresa será classificada como estando em situação iminente de insolvência.

6.4.4. Teste de Robustez

De forma a testar a robustez dos resultados obtidos nas análises anteriores, foi elaborado um teste *out-of-sample*. Assim, o objetivo foi averiguar a capacidade preditiva do modelo tendo em

¹A probabilidade do *score*: $\text{Prob}(\text{Insolvency} = 1) = \frac{1}{1 + e^{-\text{Score}}}$

conta uma subamostra da amostra inicial. O modelo final de previsão de insolvência foi aplicado a uma subamostra compreendida entre os anos 2011 e 2015 para os mesmos setores de atividade e para as mesmas regiões da NUTS2. A Tabela 29 apresenta os resultados desta estimação e o Gráfico 6 apresenta a respetiva curva ROC.

Tabela 29 - Resultados da estimação *Out-of-Sample*

Variável dependente: Insolvente	Coefficientes	Desvio Padrão	Z	p-value
Estrutura de Endividamento	0,200	0,327	0,610	0,540
Margem EBIT	1,386	0,931	1,490	0,137
Rácio Passivo – Capital Próprio	0,017	0,007	2,310	0,021
Rácio de Cobertura de Juros	0,000	0,000	-0,440	0,663
Rotação do Ativo	0,080	0,075	1,060	0,289
Dimensão	0,091	0,072	1,260	0,206
Idade	0,322	0,126	2,560	0,010
Constante	-5,884	0,703	-8,370	0,000

Observações = 13 090

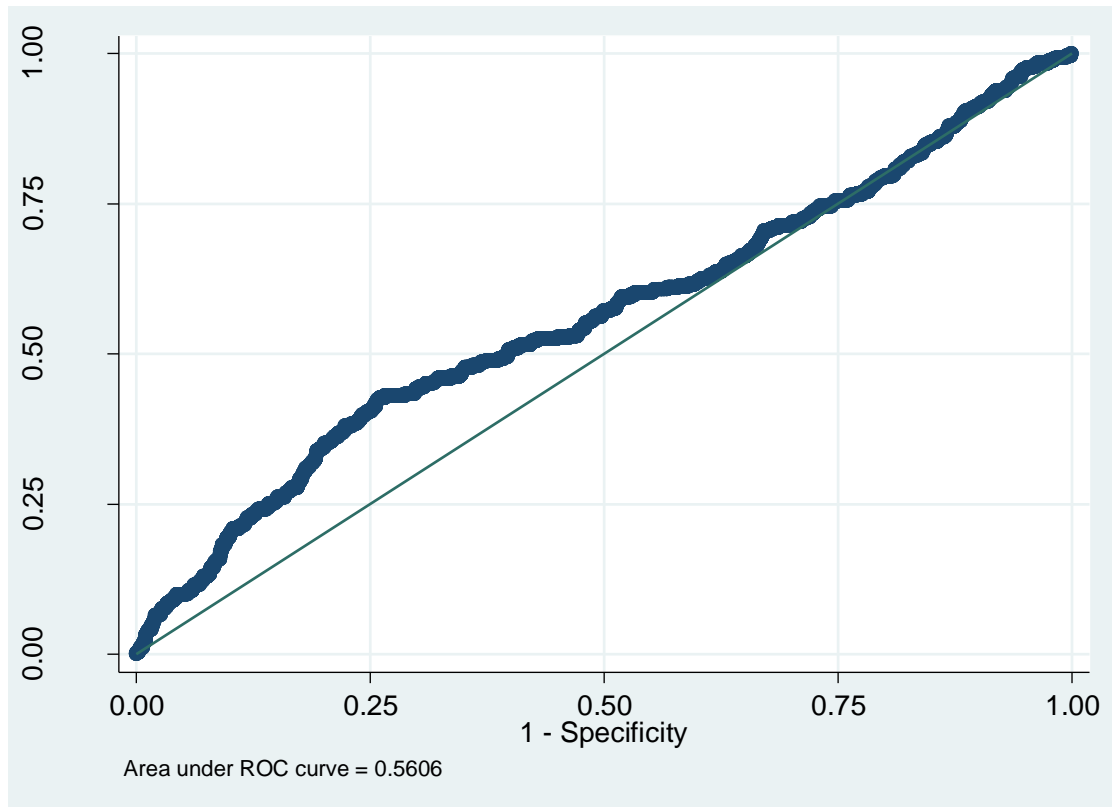
Significância Global (Prob > chi2): 0,033

Pseudo R2: 0,009

Observações Corretamente Classificadas: 60,07%

A Tabela 29 apresenta os coeficientes da estimação da equação (2) usando uma subamostra compreendida entre os anos 2011 a 2015. A variável dependente é uma *dummy* que identifica as empresas com status 1 se insolvente e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 5.2.2. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% são assinalados a negro. Os erros-padrão são *White-robust*.

Gráfico 6 - Curva ROC (*Análise Out-of-Sample*)



O Gráfico 6 apresenta a classificação das empresas em função do seu *Status* - insolvente/saudável - com base na equação (2) usando uma subamostra compreendida entre os anos 2011 a 2015.

Conforme se observa na Tabela 29 e no Gráfico 6, o teste *out-of-sample* apresenta resultados consistentes com os descritos na secção anterior. Os resultados da Tabela 29 mostram que embora alguns coeficientes tenham perdido a sua significância estatística (Estrutura de Endividamento, Margem EBIT e Rotação do Ativo), a significância conjunta do modelo ainda é evidente; o teste χ^2 é estatisticamente significativo ao nível de 5%. É também apresentada uma taxa de observações classificadas corretamente de 60,07% (no modelo anterior, esta taxa era de cerca de 65%).

A curva ROC, representada no Gráfico 6, apresenta uma precisão de 56,06% na separação entre empresas insolventes *versus* saudáveis, sendo que na análise anterior era de 57%, pelo que o diferencial entre o modelo final e a estimação *out-of-sample* é de apenas 1 ponto percentual.

6.4.5. Aplicação prática do modelo

De modo a testar a aplicação do modelo de previsão ao contexto real, o mesmo foi testado numa empresa, selecionada de forma aleatória, da nossa base de dados. Assim, a empresa com o nome fictício “XPTO”, no ano de 2016, apresenta os seguintes coeficientes:

- Estrutura de Endividamento: 0,977;
- Margem EBIT: 0,011;
- Rácio Passivo – Capital Próprio: 3,078;
- Rácio de Cobertura de Juros: 3,892;
- Rotação do Ativo: 1,222;
- Dimensão: 6,184;
- Idade: 2,485;

De seguida foi estimado o modelo:

$$Score = -5,543 + 0,355 * 0,977 + 1,562 * 0,011 + 0,016 * 3,078 - 0,001 * 3,892 + 0,100 * 1,222 + 0,065 * 6,184 + 0,218 * 2,485 = -4,068 \quad (4)$$

De forma ao *score* ser transformado na probabilidade de insolvência da empresa foi efetuado o seguinte:

$$Prob (Insolvency = 1) = \frac{1}{1 + e^{-(-4,068)}} = 0,017 \quad (5)$$

De acordo com a análise do Subcapítulo 6.4.3, o modelo final de previsão de insolvência apresenta uma zona de *cutoff* de 1,9%. Assim, como o *score* da empresa que serviu de exemplo é de -4,068, ou seja, inferior a 1,9%, a mesma é classificada como sendo uma empresa saudável.

7. Conclusão

O estudo da insolvência empresarial tem vindo a ser feito em vários países ao longo de várias décadas. A ocorrência deste evento apresenta elevada correlação com os ciclos económicos, para além de fatores como a má gestão financeira e organizacional. Tendo por base a premissa anterior e o facto de Portugal ser, quase integralmente composto por PME, ou seja, empresas que estão muito dependentes de financiamento bancário para operarem, a temática da insolvência empresarial assume particular relevância.

Desta forma, o principal objetivo deste trabalho é identificar um conjunto de variáveis consideradas determinantes na previsão de insolvência de PME do Norte e Centro de Portugal, dos diversos setores de atividade.

De acordo com o objetivo da investigação, foi selecionada uma amostra compreendida entre 2011 e 2019, incluindo 18 622 PME do Norte de Portugal e 7 633 PME do Centro de Portugal e foram construídos 16 indicadores financeiros e não financeiros, com base na literatura anterior.

A análise empírica percorreu várias etapas com o objetivo de se definir um modelo final capaz de prever a situação de insolvência.

Numa primeira fase, e como análise preliminar, foi realizada a Análise Discriminante Múltipla (ADM). Os resultados dessa análise preliminar foram comparados com os resultados da metodologia principal deste trabalho – a análise de regressão Logística. As variáveis selecionadas nesse processo foram as seguintes: Estrutura de Endividamento, Margem EBIT, Rácio Passivo-Capital Próprio, Rácio de Cobertura de Juros, Rotação do Ativo, Dimensão e Idade. Com o objetivo de mitigar problemas relacionados com a desproporcionalidade da amostra em termos de observações de empresas insolventes *versus* saudáveis, foi utilizada a metodologia *Propensity Score Matching* (PSM).

Apesar do modelo final ter sido sujeito a vários testes de robustez, a capacidade preditiva deste manteve-se estável. Os testes de robustez realizados foram os seguintes: i) controlo para efeitos temporais, de localização e de tipo de indústria; ii) estimação do modelo final com recurso a um modelo *Probit*; e iii) teste *out-of-sample* para testar a capacidade do modelo quando aplicado a amostras futuras. Nas regressões foram incluídas *dummies* para as regiões (Norte e Centro) e para os diferentes anos, de modo a controlar as características específicas de cada região e os eventos ao longo do tempo que poderiam afetar e enviesar a inferência estatística.

De realçar a capacidade preditiva do modelo final que prevê mais de 64% das observações

corretamente. Sendo que, quando se testa o modelo para dois anos antes da situação de insolvência, este prevê, aproximadamente 82% das observações corretamente e, aproximadamente, 90% para três anos antes. Assim, conclui-se que o modelo desenvolvido, encontra-se razoavelmente bem enquadrado com os resultados apresentados pela maioria dos estudos, que têm como objetivo a previsão de insolvência empresarial, cuja capacidade preditiva varia entre os 70 e os 99%.

Através desta investigação, pretende-se contribuir para o alargamento do conhecimento empírico na área das técnicas de previsão de insolvência empresarial. Uma vez que, os resultados obtidos na presente investigação não estão isentos de limitações e face à atual crise de saúde pública que mudou a forma de atuação de diversas empresas, torna-se pertinente replicar este estudo, num futuro próximo, para uma amostra mais representativa do contexto atual e com diferentes variáveis financeiras e não financeiras. Outra sugestão para os próximos estudos é a utilização de empresas *zombie*, como são o caso das empresas com custos de financiamento superiores ao EBIT ou EBITDA, nos últimos 2/3 anos, i.e., as empresas que não geram ganhos e que suportam custos de financiamento, sendo que esta classificação empresarial poderá ser uma alternativa à classificação por *Status* resultante do processo judicial da insolvência, resultante do facto de o *Status* da empresa atribuído pela ORBIS não parecer ser coincidente com o real *Status* da empresa, o que constitui uma limitação para o estudo. Outra das limitações é a disponibilidade da informação financeira que no caso de muitas empresas, é deficiente dada a informação que essas empresas disponibilizam aos outros agentes económicos (p.e, Estado e Outros Entes Públicos), uma vez que muitas das observações são perdidas por falta de informação para empresas de menor dimensão.

Por último, também pode ser relevante ajustar o modelo à relação que possa existir entre variáveis qualitativas (tais como: qualificações académicas dos recursos técnicos, a formação dos trabalhadores e a existência ou não de mudanças no tipo de produção durante a crise pandémica).

Assim sendo, e apesar de todas as limitações identificadas, este estudo contribui para o conhecimento científico dos determinantes da situação insolvência das PME do Norte e Centro de Portugal, e constitui uma ferramenta de auxílio para a gestão, principalmente em contextos de incerteza e de volatilidade constante das economias no atual contexto pandémico.

Referências bibliográficas

Abrahamse, A. & Frederikslust, R., (1975). *Discriminant analysis and prediction of corporate failure*. Rankine, Graeme (Hrsg.), European Finance Association – meetings held in 1975, Amsterdam, S. 329-373.

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

Altman, E. I. (1993). *Corporate distress diagnosis: comparing using linear discriminate analysis and neural networks*. Journal of Banking and Finance, 18, 505-529.

Altman, E. I., & Hotchkiss, E. (1993). *Corporate financial distress and bankruptcy: predict and avoid bankruptcy*. Analyze and Invest in Distressed Debt. John Wiley & Sons.

Altman, E. I., & Sabato, G. (2007). *Modelling Credit Risk for SMEs: Evidence from the U.S. Market*. Abacus, 43(3), 332-357.

Altman, E. I., Haldeman, R. G. & Narayanan, P (1977). *Zeta Analysis. A New Model to Identify Bankruptcy of Corporations*. Journal of Banking and Finance, 1, 29-54.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). *Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model*. Journal of International Financial Management & Accounting, 28(2), 131-171.

Altman, E. I., Margaine, M., Schlosser, M. & Vernimmen, P. (1974). *Statistical credit analysis in the textile industry: a French experience*. Journal of Financial and Quantitative Analysis.

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.

Beaver, W. (1966). *Financial ratios as predictors of failures. Empirical Research in Accounting: Selected Studies.*, Journal of Accounting Research, 5, 71-102.

Cherobim, A., Junior, A., & Rigo, C. (2010). *Administração financeira: princípios, fundamentos e práticas brasileiras*. Brasil: Campus.

Cochran, W. & Rubin, D. (1973), *Controlling Bias in Observational Studies: A Review*. The Indian Journal of Statistics, 35(4), 417-446.

Domenichelli, O. (2018). *Performance, financing decisions and corporate governance of Italian medium and large private family firms*. African Journal of Business Management, 12(19), 574-585.

Edmister, R. O. (1972). *Empirical test of financial ratio analysis for small business failure prediction*. Journal of Financial and Quantitative Analysis, 7(2), 1477-1493.

Fernandes, C., Peguinho, C. & Vieira, E. (2012). *Análise Financeira-Teoria e Prática*. Lisboa: Edições.

Fitzpatrick, P. (1932). *A comparison of ratios of successful industrial enterprises with those of failed companies*. The Certified Public Accountant, 727-731.

Hopp, J. C. & Leite, H.P. (1989). *O Mito da Liquidez*. Revista de Administração de Empresas, 29(4) 63-69.

Hosmer, Jr., Lemeshow, S. & Sturdivant, R. (2013). *Applied logistic regression*, 398. Hoboken, New Jersey: John Wiley & Sons.

Jackendoff, N. (1962). *A study of published industry financial and operating ratios*. Philadelphia: Temple University, Bureau of Economic and Business Research.

Jagels, M. & Coltman, M. (2004). *Hospitality management accounting*. John Wiley and Sons, Inc, 8th Edition.

Merwin, C. (1942). *Financing small corporations in five manufacturing industries*. National Bureau of Economic Research, 1926–1936.

Moreira, J. A. & Tavares, A.M. (2014). *Rentabilidade do Ativo: será indiferente a definição*

adotada? XVI Encontro AECA1, 45.

Ohlson, J. A. (1980). *Financial Ratios and the Probabilistic Prediction of Bankruptcy*. Journal of Accounting Research, 18(1), 109-131.

Pereira, H., Tavares, F. & Pacheco, L. (2015). *Determinantes da estrutura de capital das pequenas e médias empresas do vinho verde*. Revista Universo.

Platt, H. D. & Platt, M. B. (1990). *Development of a Class of Stable Predictive Variables: The Case of Bankruptcy Prediction*. Journal of Business Finance and Accounting, Spring, 31 –51.

Rodrigues, J. (2018). *O Governo e a estrutura de endividamento: O caso das PME's portuguesas*. Repositório Científico do Instituto Politécnico de Viseu.

Rosenbaum, P. & Rubin, D. (1983). *The central role of the propensity score in observational studies for causal effects*. Biometrika, 70(1) 41–55.

Sage, P. (2015). *Extension chapters on advanced techniques*. Disponível em: <https://webcache.googleusercontent.com/search?q=cache:rWq9ynQWKnEJ:https://www.researchgate.net/file.PostFileLoader.html%3Fid%3D548f502ed11b8b793c8b45bc%26assetKey%3DAS%253A273652862390274%25401442255334885+&cd=2&hl=pt-PT&ct=clnk&gl=pt>

Smith, R. & Winakor, A. (1935). *Changes in the financial structure of unsuccessful industrial corporations*. Urbana: University of Illinois.

Sousa, J. & Oliveira, I. (2014). *As variáveis de previsão da falência nas empresas portuguesas de vestuário, couro e produtos de couro*. Revista Portuguesa e Brasileira de Gestão, 13(1), 62-73.

Taffler, R. & Tisshaw, H. (1977). *Going, going, gone – four factors which predict*. Accountancy, March, 50-54.

Tamari, M. (1966). *Financial ratios as a means of forecasting bankruptcy*. Management International Review, 6(4), 15–21.

Zavgren, C. (1985). *Assessing the vulnerability to failure of American industrial firms: a Logistic Analysis*. Journal of Business Finance & Accounting, 12(1), 19–45.

Zmijweski, M. (1984). *Methodological issues related to the estimation of financial distress prediction models*. Journal of Accounting Research, 22, 59–82.

Apêndice

Tabela A - Definição das indústrias utilizadas no estudo

Indústria	Descrição	CAE
3	Indústria	10; 11; 12; 13; 14; 15; 16; 17; 18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 33
5	Abastecimento de água; esgotos, gestão de resíduos	36; 37; 38; 39
6	Construção	41; 42; 43
7	Comércio grossista e retalhista; reparação de veículos e motocicletas	45; 46; 47
8	Transporte e armazenamento	49; 50; 51; 52; 53
9	Serviços de hotelaria e alimentação	55; 56
10	Informação e comunicação	58; 59; 60; 61; 62; 63
13	Atividades profissionais, científicas e técnicas	69; 70; 71; 72; 73; 74; 75
14	Atividades administrativas e de serviço de apoio	77; 78; 79; 80; 81; 82
16	Educação	85
17	Saúde e atividades de assistência social	86; 87; 88
18	Artes, entretenimento e recreação	90; 91; 92; 93
19	Outros serviços	94; 95; 96

Esta Tabela apresenta a definição dos tipos de indústria e respetiva definição utilizados no processo de tratamento da amostra.