



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

**Os modelos preditivos do sucesso de candidaturas a fundos europeus:
o papel da manipulação de resultados**

Catarina Marçal Lagem Abrantes

Mestrado em Contabilidade

Orientador:

Professor Doutor Raul Manuel da Silva Laureano, Prof. Auxiliar, ISCTE
Business School, Departamento de Métodos Quantitativos para Gestão
e Economia

Novembro, 2020



**BUSINESS
SCHOOL**

**Os modelos preditivos do sucesso de candidaturas a fundos europeus:
o papel da manipulação de resultados**

Catarina Marçal Lagem Abrantes

Mestrado em Contabilidade

Orientador:

Professor Doutor Raul Manuel da Silva Laureano, Prof. Auxiliar, ISCTE
Business School, Departamento de Métodos Quantitativos para Gestão
e Economia

Novembro, 2020

Agradecimentos

Quero agradecer a todos os que contribuíram de alguma maneira para este projeto, e de algum modo fizeram parte deste percurso e o tornaram possível:

Ao professor Doutor Raul Laureano pela partilha de conhecimento, disponibilidade, incentivo e ajuda despendida ao longo de todo o estudo.

Ao IAPMEI, na pessoa de Pedro Cilínio por ter contribuído para que este trabalho fosse possível.

Um agradecimento especial aos meus pais, e à minha irmã por me permitirem seguir o meu caminho, mesmo que inicialmente tenha sido difícil vislumbrá-lo.

À minha família, e aos meus amigos por todo o apoio, e incentivo, mas principalmente por nunca me deixarem desistir.

À Isabel e à Joana, pela amizade que nos une, mas principalmente por serem sempre presença constante em todos os desafios, apoiando-me nos momentos mais difíceis, e acreditando sempre em mim.

À minha amiga Catarina pelo companheirismo, incentivo e sobretudo pela partilha intelectual ao longo desta caminhada.

A todos os que sempre me apoiaram, o meu sincero obrigada.

Estarei eternamente grata,

Catarina

Resumo

Portugal, desde a adesão à União Europeia até 2020, beneficiou de mais de 132 mil milhões de euros de apoio estrutural comunitário inseridos no esforço de convergência intereuropeu. No entanto, é questionada a eficácia da utilização dos fundos em Portugal, uma vez que muitas empresas apoiadas ficam aquém dos objetivos esperados. Assim, importa identificar as causas para este insucesso, concretizado no estudo do impacto desta ineficiência quer na elegibilidade da candidatura, quer na avaliação do projeto. Neste sentido, este estudo avalia o impacto da manipulação de resultados na previsão do sucesso de candidaturas na atribuição de fundos europeus.

Esta investigação foi suportada por uma revisão sistemática da literatura, concretizada em 36 artigos, e por técnicas de análise de dados, nomeadamente regressões lineares e árvores de decisão, com os algoritmos CART e CHAID. Para tal, a partir das demonstrações financeiras de 2010 a 2020, de 14.122 empresas, o que corresponde a 22.422 candidaturas, foram aplicados quatro modelos baseados em *accruals*, assim como foram calculadas diversas variáveis independentes necessárias aos modelos preditores da elegibilidade e do sucesso.

Os resultados permitiram identificar uma relação entre a manipulação de resultados e o sucesso dos projetos, prevendo uma percentagem de exemplos corretamente classificados de 84%. Como conseguinte, o principal contributo deste estudo é ajudar as instituições portuguesas a tornar a atribuição dos fundos mais eficiente e eficaz, de forma a que estes tenham um impacto maior, e positivo, no desenvolvimento da economia portuguesa.

Palavras-chave: Manipulação de resultados, *Accruals*, Fundos Europeus, Árvores de Decisão

JEL Classification System: M10, M41

Abstract

Portugal, since joining the European Union until 2020, has benefited from more than 132 billion euros of Community structural support as part of the inter-European convergence effort. However, the effectiveness of the use of funds in Portugal is questioned as many supported companies fall short of their expected goals. Thus, it is important to identify causes for this failure, i.e., to understand its impact on both application eligibility and project evaluation. In this sense, this study seeks to evaluate the impact of the manipulation of results in predicting the success of applications in the allocation of European funds.

This research was supported by a systematic review of the literature, carried out in 36 articles, and by data analysis techniques, namely linear regressions and decision trees, with the CART and CHAID algorithms. To this end, from the 2010 to 2020 financial statements of 14,122 companies, corresponding to 22,422 applications, four accrual-based models were applied, as well as several independent variables required for predictor models of eligibility and success.

The results identified a relationship between the earnings management and the success of the projects, achieving an accuracy of 84%. Therefore, the main contribution of this study is to help Portuguese institutions to make the allocation of funds more efficient and effective, so that they have a greater and positive impact on the development of the Portuguese economy.

Keywords: Earnings Management; Accruals; European Funds; Decision Trees

Classification JEL: M10, M41

Índice Geral

Resumo.....	iii
Abstract	v
Índice Geral	vii
Índice de Figuras	ix
Índice de Tabelas.....	xi
Lista de Acrónimos e Siglas	xiii
1. Introdução	1
1.1 Tema e sua importância.....	1
1.2 Problema e questão de investigação.....	3
1.3 Objetivos e contributos	6
1.4 Abordagem Metodológica	6
1.5 Estrutura e organização da dissertação.....	7
2. Revisão de Literatura	9
2.1 Protocolo para a revisão sistemática.....	9
2.2 Métodos para a deteção de MR	13
2.2.1 Análise de rácios e tendências.....	14
2.2.2. Observação e análise da distribuição dos resultados.....	14
2.3. Modelos baseados <i>accruals</i>	16
2.3.1. Tipologias de modelos	16
2.3.3. Contextos de aplicação	23
2.3.2. Avaliação dos resultados	26
2.4 Avaliação dos artigos científicos relevantes	28
3. Metodologia.....	31
3.1 Posicionamento da Investigação	31
3.2 Amostra e dados	31
3.2.1 Processo de seleção da amostra.....	31
3.2.2 Caracterização da amostra	32
3.3 Caracterização e mensuração dos dados a analisar	33
3.3.1 Variáveis dependentes adotadas – Elegibilidade; Avaliação do projeto.....	33
3.3.2 Variáveis independentes adotadas	34
3.4 Técnicas de análise de dados.....	37
3.5 Modelo com Árvores de Decisão	38
4. Resultados e Discussão	43
4.1 Avaliação da manipulação de resultados nas candidaturas	43

4.2	Relação entre a manipulação de resultados e a elegibilidade das candidaturas.....	46
4.3	Modelo preditivo para elegibilidade das candidaturas	49
4.4	Relação entre a manipulação de resultados e a avaliação do projeto	53
4.5	Modelo preditivo para a avaliação do projeto	56
5.	Conclusões	61
5.1	Contributos	62
5.2	Limitações e pistas futuras de investigação	63
	Referências Bibliográficas	65

Índice de Figuras

Figura 1: Fases do processo de atribuição de incentivos, percorridas ao longo dos anos.....	4
Figura 2: Processo de seleção dos artigos para revisão	10
Figura 3: Nuvem de palavras-chave dos artigos selecionados	13
Figura 4: Importância das variáveis preditoras da elegibilidade, relativas ao modelo H.....	50
Figura 5: Distribuição da elegibilidade prevista para as três variáveis mais importantes	51
Figura 6: Importância das variáveis preditoras da avaliação do projeto, relativas ao modelo I.57	
Figura 7: Distribuição do sucesso prevista para as três variáveis mais importantes	58

Índice de Tabelas

Tabela 1: Critérios de inclusão e de exclusão.....	10
Tabela 2: Artigos incluídos na revisão sistemática entre 2019 a 2015	11
Tabela 3: Artigos incluídos na revisão sistemática entre 2014 a 2010	11
Tabela 4: Artigos incluídos na revisão sistemática entre 2009 a 1998	12
Tabela 5: Critério de qualidade para avaliação dos artigos.....	12
Tabela 6: Variáveis constituintes do M-Score de Beneish.....	20
Tabela 7: Tipologias de modelos utilizadas nos artigos da RSL.....	22
Tabela 8: Contextualização da aplicação da MR dos artigos da RSL	24
Tabela 9: Avaliação dos resultados dos artigos alvo da RSL.....	27
Tabela 10: Avaliação da qualidade dos artigos que integram a RSL	28
Tabela 11: Processo de seleção da amostra.....	32
Tabela 12: Distribuição de candidaturas por setor de atividade entre 2012 e 2019	32
Tabela 13: Distribuição de candidaturas por NUTS II e dimensão entre 2012 e 2019.....	33
Tabela 14: Distribuição das candidaturas por elegibilidade entre os anos 2012 e 2019.....	34
Tabela 15: Distribuição das candidaturas por avaliação do projeto entre os anos 2012 e 2019.....	34
Tabela 16: Síntese dos Modelos de <i>Accruals</i> Não Discricionários – Regressão OLS	35
Tabela 17: Alterações às variáveis do modelo de Beneish.....	35
Tabela 18: Descrição das variáveis independentes	36
Tabela 19: Principais técnicas de análise de dados utilizadas.....	38
Tabela 20: Parametrização dos modelos preditivos selecionados para a elegibilidade	39
Tabela 21: Parâmetros dos modelos selecionados para a avaliação do projeto	40
Tabela 22: Matriz de classificação	40
Tabela 23: Métricas de avaliação de qualidade dos modelos de classificação	41
Tabela 24: Medidas descritivas das variáveis de manipulação de resultados	44
Tabela 25: Matriz de correlações de Pearson na manipulação de resultados.....	45
Tabela 26: Distribuição do Kasznick por elegibilidade da candidatura	47
Tabela 27: Distribuição do Kothari por elegibilidade da candidatura	47
Tabela 28: Distribuição do McNichols por elegibilidade da candidatura.....	48
Tabela 29: Distribuição do M-Score por elegibilidade da candidatura	48
Tabela 30: Resultados dos modelos preditivos para a elegibilidade.....	49
Tabela 31: Importância das variáveis dos modelos preditivos para a elegibilidade	50
Tabela 32: Distribuição do Kasznick por avaliação do projeto	53
Tabela 33: Distribuição do Kothari por avaliação do projeto.....	54
Tabela 34: Distribuição do McNichols por avaliação do projeto.....	54

Tabela 35: Distribuição do M-Score por avaliação do projeto.....	55
Tabela 36: Resultados dos modelos preditivos para a avaliação do projeto	56
Tabela 37: Importância das variáveis dos modelos preditivos para a avaliação do projeto	56

Lista de Acrónimos e Siglas

AICEP – Agência para o Investimento e Comércio Externo de Portugal

BvD – *Bureau van Dijk*

CAE-Rev.3 – Classificação Portuguesa de Atividades Económicas, Revisão 3

CART – *Classification and Regression Trees*

CHAID – *Chi-squared Automatic Interaction Detection*

IAPMEI – Agência para a Competitividade e Inovação

IES – Informação Empresarial Simplificada

MBA – Modelos Baseados em *Accruals*

MR – Manipulação de resultados

PIB – Produto Interno Bruto

PME – Pequena e Média Empresa

RL – Revisão da Literatura

RSL – Revisão Sistemática da Literatura

SNC – Sistema de Normalização Contabilística

UE – União Europeia

1. Introdução

A manipulação de resultados nos projetos candidatos a fundos europeus, em particular, a sua deteção, é uma área que os contabilistas devem dar especial atenção. Esta introdução visa, em concreto, apresentar o conceito de manipulação de resultados (MR) e realçar a sua importância e respetivos incentivos, bem como expor o problema e os objetivos desta investigação. É abordada ainda, de que de uma forma breve, a metodologia bem como a estrutura do presente estudo.

1.1 Tema e sua importância

São várias as definições que existem para MR, não existindo, por isso, uma definição consensual para descrever esta prática (Dechow *et al.*, 1996). A MR é uma questão complexa, identificada e interpretada em estudos anteriores, em função das diferentes perceções dos investigadores e onde a sua inconsistência é espelhada pelo seu carácter ambíguo e pouco mesurável (Callao *et al.*, 2010).

Em Portugal, num sentido mais amplo e generalizado, MR é frequentemente associada a gestão de resultados (do inglês *Earnings Management*), uma vez que a amostra utilizada em cada investigação é composta por empresas de pequena dimensão, dificultando ao investigador a distinção entre gestão de resultados e prática fraudulenta (Dias, 2015).

Perante esta complexidade na definição de MR, e até de confusão entre expressões, para esta investigação considera-se MR como um processo onde deliberadamente, e de acordo com os princípios contabilísticos, fazem-se escolhas de modo a atingir um determinado nível de resultados (Davinson *et al.*, 1987) ou altera-se a perceção do desempenho da empresa através da informação financeira reportada aos seus utilizadores (Healy & Wahlen, 1999)¹.

No entanto, outros autores desviam-se ligeiramente desta definição. Por exemplo, Shipper (1989) realça que tem que se tratar de uma intervenção pura e profunda do processo de relato externo, com o objetivo de obter um determinado benefício privado, contrária aquela que deveria ser uma intervenção de carácter imparcial, ao qual Mulford e Comiskey (2011) acrescentam a necessidade de agir em direção a um objetivo pré-determinado, imposto pela gestão ou para alisamento de resultados.

Por fim, importa referir Beneish (2001), um dos autores que mais se tem dedicado ao tema, que identifica duas perspetivas sobre MR. Por um lado, a vertente oportunista, na qual os gestores têm o objetivo claro de adulterar as demonstrações financeiras em prol de si mesmos. Por outro, a distorção informativa, onde os gestores alteram as demonstrações financeiras de modo a transmitirem as suas perspetivas privadas sobre o desempenho futuro da empresa, com intenção de gerar conteúdo

¹ Uma vez que na literatura são identificadas várias expressões para abordar o tema da MR nas demonstrações financeiras, neste estudo consideram-se as expressões suprarreferidas como sinónimos, fazendo apenas utilização da expressão de MR.

informativo de maior qualidade. No entanto, e de acordo com o mesmo autor, muita da investigação realizada sobre a MR tem-se debruçado sobre a perspetiva oportunista, tendo então a vertente informativa sido pouco testada na literatura.

Atendendo às implicações da MR, a sua deteção é muito importante. A crescente globalização do mercado empresarial e interdependência das economias ampliaram a urgência de uma linguagem comum que possa ser compreendida pela generalidade dos agentes económicos. O relato financeiro constitui em si o veículo primordial do fluxo de informação entre empresas e a sua envolvente externa, onde se incluem as demonstrações financeiras.

As demonstrações financeiras têm como principal objetivo proporcionar, aos variados agentes económicos, informação útil acerca da posição financeira, desempenho financeiro e fluxos de caixa da entidade (Lourenço *et al.*, 2015), e de forma a "...que seja útil a um vasto leque de utentes na tomada de decisões económicas."²

Assim, a utilidade da informação financeira encontra-se intrinsecamente relacionada com a qualidade da mesma, uma vez que as decisões tomadas dependem da informação na qual assentam. A estrutura conceptual do Sistema de Normalização Contabilística (SNC) acrescenta, nos seus parágrafos 24 a 46, que a utilidade da informação financeira é determinada por três características qualitativas: relevância, fiabilidade e comparabilidade². Estas características, juntamente com conceitos, princípio e normas contabilísticas adequadas devem promover uma imagem verdadeira e apropriada da posição financeira e do resultado das operações das empresas.

Contudo, Dias (2015) alerta para a impossibilidade da conjugação simultânea e de forma igualitária das três características qualitativas anteriormente referidas, podendo resultar na exaltação de umas em detrimento de outras. Neste sentido, sugere que as demonstrações, para além de irem ao encontro dos seus objetivos, devem privilegiar um equilíbrio entre a relevância e a fiabilidade da informação que é prestada aos utentes.

Para a subjetividade das demonstrações financeiras contribui, ainda, a flexibilidade patente nas políticas de contabilização, que permite ao gestor a adoção de critérios que no seu entender vão ao encontro com os objetivos da empresa (e.g. critérios de mensuração). Este facto, mesmo não infringindo a legalidade, pode levar ao relato de informação que espelha uma realidade que não existe. No entanto, esta não deverá constituir um processo de fraude.

De forma a garantir uma maior objetividade das demonstrações financeiras, tendo como objetivo o aumento da qualidade da informação, muitos autores sugerem que as normas contabilísticas devem ser claras quanto ao método de contabilização. No entanto, Healy e Wahlen (1999) contrapõem que,

² Estrutura Conceptual do SNC, Aviso n.º 8254/2015, 29 Julho de 2015 – Série II – Nº 146.

se a flexibilidade ilimitada não é desejada, a sua total eliminação nas normas também não constitui uma solução ótima devido à complexidade de previsão do sistema.

Como referido, a MR corresponde a um processo intencional, onde se exploram os princípios contabilísticos geralmente aceites (Baralexix, 2004), uma vez que existem incentivos para que se reportem resultados dissemelhantes àqueles que resultariam de um processo neutro, ou seja, sem a existência de intervenção ativa da gestão (Beneish, 2001). De facto, esta situação leva a que os gestores de uma determinada organização enfrentem um dilema entre a maximização do lucro, que visa a satisfação dos acionistas, e a maximização das respetivas ambições pessoais, quer a nível remuneratório, quer ao nível de progressão ou manutenção do seu cargo dentro da organização (Dias, 2015).

De facto, e de acordo com a teoria positivista da contabilidade formulada por Watts e Zimmerman (1986), os incentivos e as motivações à adoção de práticas de MR centram-se em três pilares: indexação da remuneração dos gestores; nível de endividamento das empresas; e custos políticos. A estes, Mulford e Comiskey (2011), acrescentam ainda, por um lado, os efeitos sobre compensação de executivos, efeitos sobre a cotação das ações, efeitos sobre o custo e o acesso ao endividamento e, por outro, os efeitos políticos ou de regulação. Baralexix (2004) acrescenta o incentivo relacionado com a obtenção de financiamento externo (entre elas, as instituições financeiras), onde é necessário que as organizações cumpram determinados requisitos para que seja possível a concessão de crédito, nomeadamente resultados elevados. No entanto, Moreira (2008) distancia-se dos autores anteriores, destacando o incentivo relacionado com o imposto sobre o rendimento, em linha com o preconizado por Boynton *et al.* (1992) e Guenther (1994) relativamente a alterações de taxa de imposto.

Numa outra vertente, a dos incentivos europeus, uma vez levantada dúvidas sobre a eficácia da utilização dos fundos em Portugal, Fernandes e Laureano (2019) sugerem que o acesso aos incentivos europeus, da União Europeia (UE), podem constituir um incentivo à MR, uma vez que as empresas candidatas são obrigadas a cumprir determinados requisitos, que por vezes, estão longe da sua realidade. De facto, os incentivos financeiros da UE, em geral, assumem a forma de financiamento externo (reembolsável ou não), estando, assim, estes autores alinhados com o preconizado por Baralexix (2004).

1.2 Problema e questão de investigação

Nos últimos anos, tem-se vindo a observar uma tendência global de crescimento, quer dos níveis de fraude, quer do impacto que esta crescente onda de crimes económicos tem nas organizações (ACFE, 2018).

Desde a adesão de Portugal à UE até ao ano de 2020, através dos sucessivos quadros comunitários de apoio, Portugal beneficiou de mais de 132 mil milhões de euros, num período de 35 anos, de apoio

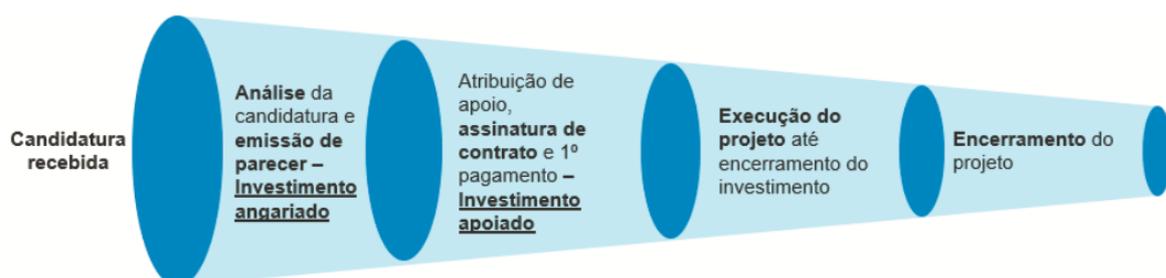
estrutural comunitário, o que representa entre dois a três por cento do Produto Interno Bruto (PIB) português ao longo destes anos (Pires, 2017). No entanto, Stec e Grzebyk (2018) colocam em causa a eficácia da utilização dos fundos, uma vez que Portugal se posiciona entre os países com menor índice de implementação das estratégias da Europa 2020. Adicionalmente, Fernandes e Laureano (2019) sugerem que a falta de eficácia na atribuição dos fundos centra-se no facto de as candidaturas a estes programas requererem o cumprimento de determinados requisitos. Por vezes, as empresas não conseguem atingir nos prazos destinados, levando à suspeição que estas possam ter a tentação de manipular os resultados (OLAF, 2015; ECA, 2017), de forma a cumprirem todos os requisitos e assim beneficiarem dos fundos.

Para evitar o mau uso dos dinheiros públicos europeus, a Comissão Europeia desenvolveu uma estratégia operacional de combate à fraude e um sistema de deteção precoce e de exclusão, que com o apoio de um painel interinstitucional decide sobre a exclusão de operadores económicos do financiamento da UE, devido a fraude ou corrupção (OLAF, 2018). No entanto, estudos recentes, publicados pelo organismo responsável por controlar a utilização dos fundos da UE, conclui que a Comissão Europeia não possui informação abrangente sobre a dimensão, a natureza e as causas da fraude, para além de nunca ter efetuado qualquer avaliação relativa à fraude não detetada (OLAF, 2018).

Em Portugal, os organismos intermédios de programas, com destaque para a AICEP (Agência para o Investimento e Comércio Externo de Portugal) e IAPMEI (Agência para a Competitividade e Inovação), têm como tarefas a análise dos processos de candidatura, a decisão de atribuição de apoio financeiro e a verificação dos incentivos, como também a identificação de situações irregulares que tem origem na MR, e declarações para efeitos das candidaturas (OLAF, 2015; ECA, 2017).

Na Figura 1 apresenta-se o processo de atribuição de incentivos financeiros, destacando-se as fases da análise da candidatura com decisão de atribuição de apoio, da execução do projeto, nomeadamente até atingir o ano de cruzeiro, e, finalmente, a última fase do encerramento. Neste estudo, são alvo de avaliação o momento em que a candidatura é elegida, e o momento em que o projeto termina e é avaliado, mas com base na informação disponível aquando do recebimento e da análise da candidatura.

Figura 1: Fases do processo de atribuição de incentivos, percorridas ao longo dos anos



Fonte: AICEP (2018:1)

Neste contexto, é importante detetar a manipulação (que pode, nas situações mais graves, ser fraude) nas declarações financeiras sujeitas a análise por parte destes organismos. Para este fim, existem diferentes métodos, sendo o mais difundido na literatura o método baseado em *accruals*, onde estes resultam do desencontro entre o momento do fluxo de caixa e o momento do reconhecimento do resultado operacional (Callao *et al.*, 2010) e que têm sido utilizados com sucesso na identificação de MR (Im & Nam, 2019; Lin & Liao, 2015).

Assim, e no âmbito do tema da MR, a presente investigação tem como primeira finalidade gerar conhecimento que permita aumentar a eficiência e a eficácia na atribuição dos fundos europeus às PME, de forma a que apenas as empresas cumpridoras acedam aos mesmo. Neste sentido, esta investigação visa responder à seguinte questão de investigação: como se pode detetar a manipulação de resultados no processo das candidaturas a fundos europeus e qual o seu impacto no sucesso das mesmas?

O sucesso dos programas de atribuição de fundos europeus coloca em causa tanto as candidaturas, momento anterior à elegibilidade da empresa, como o projeto em si, ou seja, após a empresa ser elegida para o programa (investimento angariado). Deste modo, e em consonância com a questão de investigação, formulam-se duas hipóteses de investigação que procuram espelhar esses dois momentos:

- Hipótese 1: A manipulação de resultados influencia o sucesso da candidatura a incentivos financeiros, esperando-se diferentes níveis de manipulação entre candidaturas elegíveis e não elegíveis.

Pinheiro (2008) verificou que as empresas que se candidatam a subsídios de investimento, e que não cumprem os requisitos necessários, observam-se um ajustamento significativo da informação financeira no período pré-candidatura, o que sugere a existência de MR. No entanto, numa perspetiva similar, testando os incentivos contratuais, DeFond e Jiambalvo (1994), e Sweeney (1994) comprovam que empresas próximas da violação das cláusulas dos contratos de dívida apresentam tendência para alterar as suas políticas contabilísticas. Analogamente, prevê-se que empresas com maior tendência em alterar a sua informação financeira, com vista em atingir os requisitos necessários, tenham maior probabilidade de ser elegidas para o programa de incentivos financeiros.

- Hipótese 2: A manipulação de resultados influencia a avaliação final dos projetos de investimento aprovados, esperando-se diferentes níveis de manipulação consoante o projeto tenha sucesso ou insucesso.

Noutra vertente, e desta vez mais a jusante no processo de candidaturas a fundos, o deficitário aproveitamento do apoio estrutural prende-se com o sucesso dos projetos aprovados, ou seja, as empresas que MR têm uma menor propensão para atingir o sucesso do projeto, se anteriormente alteraram a sua informação financeira (Fernandes & Laureano, 2019). Neste sentido, é expectável que

as empresas que mais fielmente espelham os seus resultados, à data da candidatura, tenham maior probabilidade de atingir o sucesso da mesma.

1.3 Objetivos e contributos

Atendendo à questão de investigação formulada, esta investigação tem como principais objetivos: O1) Avaliar o impacto da manipulação de resultados no sucesso de candidaturas a incentivos financeiros, e O2) Avaliar o impacto da manipulação de resultados no sucesso dos projetos financiados.

De forma a concretizar estes dois objetivos principais, são definidos quatro objetivos específicos para cada um dos principais:

O1.1: Avaliar o grau de manipulação de resultados nas candidaturas recorrendo a diferentes modelos discricionários de *accruals*;

O1.2: Caracterizar a relação entre manipulação de resultados e a elegibilidade das candidaturas;

O1.3: Criar modelo preditivo da elegibilidade das candidaturas;

O2.1: Caracterizar a relação entre a manipulação de resultados e a avaliação do projeto aprovado;

O2.2: Criar modelo preditivo da avaliação do projeto aprovado.

A concretização dos objetivos e, conseqüentemente, a resposta à questão de investigação, vem contribuir, por um lado, para a literatura ao colmatar uma lacuna no conhecimento, já que os modelos de *accruals* têm sido aplicados em inúmeras situações de suspeita de MR no relato financeiro, mas nunca em empresas potencialmente candidatas e dos processos de candidatura quando estas formalizam as suas candidaturas. Por outro, para os profissionais do IAPMEI, este estudo permite sinalizar a existência de manipulação e identificar perfis de empresas propensas a essa manipulação e, assim, vem contribuir para uma atribuição dos fundos mais eficiente e eficaz (e até justa) e, conseqüentemente, com maior impacto no desenvolvimento do país.

1.4 Abordagem Metodológica

Atendendo aos objetivos, o estudo assume cariz quantitativo, uma vez que é baseada em evidência empírica (Major, 2009).

Assim, este estudo tem por base uma amostra de 34 mil candidaturas a fundos europeus, indicadas pelo IAPMEI. A evidência a recolher, isto é, os dados a recolher são provenientes, essencialmente, de duas fontes. Por um lado, o IAPMEI disponibiliza toda a informação que permite caracterizar o projeto candidato e seu sucesso, e, também, a empresa candidata ao programa de incentivos financeiros. Por outro, recorre-se à base de dados AMADEUS para recolher os dados que permitem caracterizar a situação financeira em diversos momentos (antes da candidatura, no ano da candidatura e no ano da avaliação do projeto). A integração dos dados é possível através do NIF da empresa.

As estratégias para assegurar a qualidade da investigação assentam, essencialmente, em dois pilares. Primeiro, toda a investigação é acompanhada por especialistas, quer em fundos europeus, quer em técnicas de deteção de fraude e MR. Em segundo, os métodos a utilizar na análise da evidência recolhida, após uma rigorosa avaliação da qualidade dos dados, são métodos amplamente utilizados e que já demonstraram produzir bons resultados.

De facto, os dados são analisados recorrendo a diferentes técnicas estatísticas. Inicialmente, recorre-se a técnicas de estatística descritiva uni e bivariada e a testes de hipóteses para avaliar a significância das relações entre variáveis encontradas (Laureano, 2020). Depois, atendendo à complexidade do problema, recorre-se a técnicas de *data mining*, designadamente a árvores de decisão, para a criação dos modelos preditivos da elegibilidade e do sucesso das candidaturas.

Por fim, todas estas análises são suportadas por uma revisão sistemática da literatura (RSL) sobre modelos de MR baseados em *accruals*, seguindo um rigoroso protocolo de RSL, que permitem concretizar os objetivos estabelecidos e, conseqüentemente, responder à questão de investigação formulada (Vilelas, 2017).

1.5 Estrutura e organização da dissertação

De forma a compreender de que modo a MR impacta a afetação dos fundos europeus, desenvolvemos este trabalho organizando-o em cinco capítulos que refletem as diferentes fases da investigação.

Neste primeiro capítulo são abordados o tema e a questão de investigação, apresentados os objetivos e os principais contributos e, ainda, descreve-se a metodologia adotada. O segundo capítulo reflete o enquadramento teórico, designado por revisão da literatura, onde é apresentado o conceito de métodos baseados em *accruals* para a deteção da MR. Adicionalmente, é apresentado o protocolo da RSL, onde se destacam quatro dimensões analisadas, entre elas a qualidade dos artigos. A metodologia utilizada nesta dissertação é apresentada no capítulo três, onde é descrito todo o processo metodológico de recolha e tratamento de dados bem como as técnicas estatísticas utilizadas na análise dos mesmos. No quarto capítulo são apresentados e discutidos os resultados obtidos. Por fim, no quinto capítulo, apresentam-se as conclusões e identificam-se os contributos desta investigação. No seu final identificam-se as limitações e apresentam-se pistas futuras de investigação.

2. Revisão de Literatura

Este capítulo aborda a MR e, em concreto, os modelos baseados em *accruals* para a sua deteção. Primeiramente, é apresentado o protocolo da RSL, bem como todo o processo até obter os artigos analisados. Em segundo, tendo como ponto de partida os métodos mais comuns para a deteção de MR, analisam-se os conceitos, vantagens e limitações de cada um. No terceiro ponto, tendo por base uma RSL, são abordados os diversos modelos de *accruals*, onde se exploram a tipologia, o contexto de aplicação, e a avaliação da qualidade dos mesmos. Por fim, os artigos alvo da revisão sistemática são avaliados em termos da sua qualidade recorrendo a diferentes critérios de avaliação.

2.1 Protocolo para a revisão sistemática

Revisão sistemática da literatura trata-se de um método de pesquisa orientado por protocolos específicos, destacando-se o carácter de reprodutibilidade do estudo (Caldas *et al.*, 2017). O protocolo deve apresentar de forma explícita, as bases de dados bibliográficas consultadas, a *query* utilizada, o processo de seleção dos artigos, incluindo os critérios de inclusão e exclusão dos artigos, bem como a avaliação da qualidade dos mesmos (Vilelas, 2017).

A revisão de literatura (RL) do presente estudo, tem como base uma RSL, onde o seu objetivo é elaborar o estado da arte em MR com base em métodos empíricos baseados em *accruals*. Atendendo a este objetivo, esta revisão propõe-se a responder à pergunta: *como se podem aplicar os modelos baseados em accruals ao estudo da manipulação de resultados nas candidaturas aos fundos europeus, que é uma forma de financiamento?* Mais especificamente, responde às cinco questões seguintes: (1) Como se caracterizam os métodos empíricos baseados em *accruals*?; (2) Quais os modelos utilizados para de deteção de manipulação de resultados nas declarações financeiras?; (3) Em que contextos são utilizados (perfis de empresas – amostra e para que fins) os diferentes modelos?; (4) Como são avaliados os resultados obtidos?.

Os artigos incluídos nesta RL foram alvo de uma seleção primária, definida de acordo com as fontes de estudos anteriores, palavras-chave e com base em critérios definidos de inclusão e exclusão. A base de dados científica utilizada foi a *web of science* (www.webofknowledge.com), tendo sido selecionada devido à sua abrangência e relevância, académica e prática, dos *journals* indexados e por incluir publicações relevantes nas áreas científicas da investigação, e por ser amplamente utilizada em RL (Vilelas, 2017).

A pesquisa automática de artigos na base de dados, através de uma *query* aplicada ao tópico do artigo, isto é, a procura de palavras/expressões efetua-se nos campos título, resumo, autor e palavras-chave. As palavras-chave utilizadas na *query* resultam da identificação na literatura de sinónimos de MR, *accruals* e demonstrações financeiras, e que posteriormente foram alvo de avaliação por parte de especialistas. A *query* definida foi: "(Discretionary Accrual*" or "Non Discretionary

Accrual*" or "Manipulation of Accrual*" or "Accrual* Manipulation*") AND ("Fraud Detection" or "Anomal* Detection" or "Fraud*" or "Detection" or "Anomal*" or "Abnormal*") AND ("Earning* Management" or "Earning* Manipulation*" or "Accounting Irregular*" or "Result* Manipulation*" or "Revenue* Manipulation*" or "Earning* Quality")) AND ("Financial Statement*" or "Financial Report*" or "Account* Statement*" or "Report*")". Deste processo preliminar de pesquisa resultaram 69 artigos, que posteriormente foram sujeitos aos critérios de inclusão e de exclusão, apresentados na Tabela 1.

Tabela 1: Critérios de inclusão e de exclusão

Critérios de Inclusão	Artigos que abordam MR;
	Artigos que analisam as contas (incluindo indicadores financeiros) das empresas;
	Artigos que abordam o método empírico baseado em <i>accruals</i> .
Critérios de Exclusão	Posters, <i>papers</i> , tutoriais e editoriais;
	Artigos duplicados;
	Artigos que não estão em língua inglesa;
	Artigos que não tenham fator de impacto;
	Artigos que não estejam totalmente disponíveis;
	Artigos que não tenham componente empírica ou com lacunas metodológicas.

A aplicação dos critérios de inclusão e exclusão obrigou à leitura do resumo, da introdução e da metodologia, levando à exclusão de 33 artigos, estando o processo resumido na Figura 2. Desta forma, foram incluídos na RL desta investigação 36 artigos, apresentados nas Tabela 2, Tabela 3 e Tabela 4.

Figura 2: Processo de seleção dos artigos para revisão

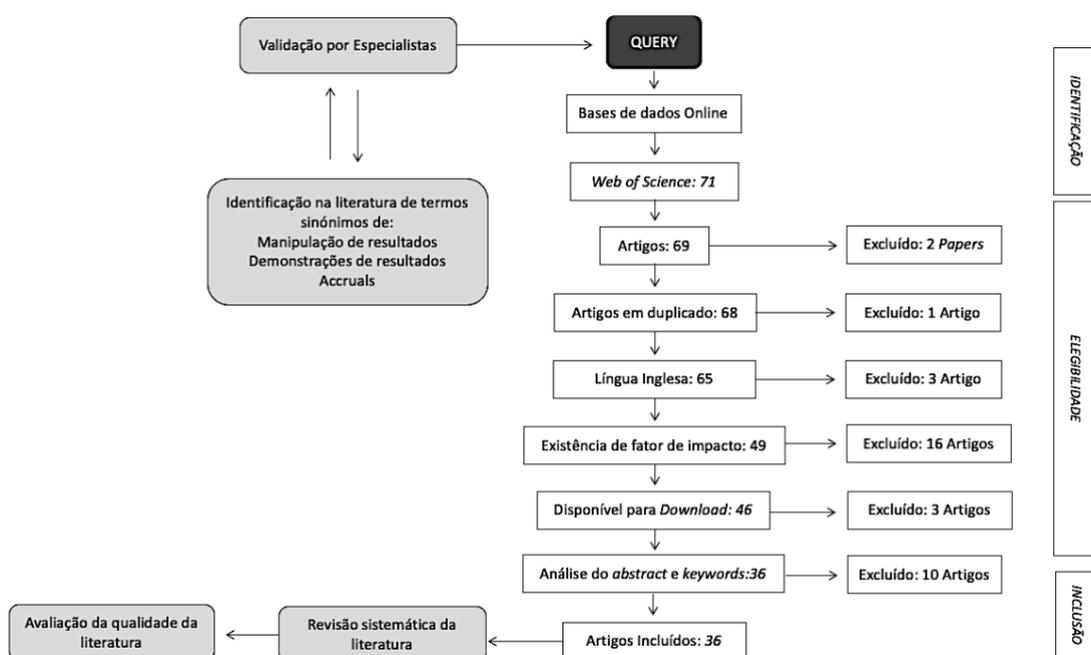


Tabela 2: Artigos incluídos na revisão sistemática entre 2019 a 2015

ID	Ano	Título	Autores	Journal	Quartil
1	2019	Does Ethical Behavior of Management Influence Financial Reporting Quality?	Im, C.; Nam, G.	Sustainability	Q2
2	2019	Earnings Management in Chapter 11 Bankruptcy	Fisher, T. C. G.; Gavious, I.; Martel, J.	Abacus-A Journal of Accounting Finance and Business Studies	Q2
3	2019	Financial Reports and Social Capital	Jha, A.	Journal of Business Ethics	Q1
4	2018	Opportunistic financial reporting around municipal bond issues	Beck, A. W.	Review of Accounting Studies	Q2
5	2018	Managerial reporting behavior around exchange switching: Consideration of current and future performance	Lin, W.; Liao, T.	International Review of Economics & Finance	Q2
6	2018	Does there prevail momentum in earnings management for seasoned equity offering firms?	Chang, C.; Lin, H. W.	International Review of Economics & Finance	Q2
7	2018	The impact of auditor conservatism on accruals and going concern opinion: Iranian angle	Salehi, M.; Tarighi, H.; Sahebkar, H.	International Journal of Islamic and Middle Eastern Finance and Management	Q4
8	2018	How Do Financial Constraints Relate to Financial Reporting Quality? Evidence from Seasoned Equity Offerings	Kurt, A. C.	European Accounting Review	Q1
9	2018	Accrual management as an indication of money laundering through legally registered Mafia firms in Italy	Ravenda, D.; Valencia-Silva, M. M.; Maria Argiles-Bosch, J.; Garcia-Blandon, J.	Accounting Auditing & Accountability Journal	Q1
10	2017	Managerial Discretion in Accruals and Informational Efficiency	Perotti, P.; Windisch, D.	Journal of Business Finance & Accounting	Q2
11	2017	Do social ties between individual auditors and client CEOs/CFOs matter to audit quality?	Qi, B.; Yang, R.; Tian, G.	Asia-Pacific Journal of Accounting & Economics	Q4
12	2016	Problem directors on the audit committee and financial reporting quality	Habib, A.; Bhuiyan, Md. B. U.	Accounting and Business Research	Q2
13	2015	External investigations and disciplinary sanctions against auditors: the impact on audit quality	De Fuentes, C.; Illueca, M.; Consuelo P., M.	Series-Journal of the Spanish Economic Association	Q3
14	2015	Exchange listing type and firm financial reporting behavior	Lin, W.; Liao, T.	International Review of Economics & Finance	Q3
15	2015	Managing Discretionary Accruals and Book-Tax Differences in Anticipation of Tax Rate Increases: Evidence from China	Wong, R. M. K.; Lo, A. W. Y.; Firth, M.	Journal of International Financial Management & Accounting	Q2

Tabela 3: Artigos incluídos na revisão sistemática entre 2014 a 2010

ID	Ano	Título	Autores	Journal	Quartil
16	2013	Does Investor Sentiment Affect Earnings Management?	Simpson, A.	Journal of Business Finance & Accounting	Q2
17	2013	Female Auditors and Accruals Quality	Ittonen, K.; Vahamaa, E.; Vahamaa, S.	Accounting Horizons	Q3
18	2012	Do UK firms manage earnings to meet dividend thresholds?	Atieh, A.; Hussain, S.	Accounting and Business Research	Q2
19	2011	Detecting and Predicting Accounting Irregularities: A Comparison of Commercial and Academic Risk Measures	Price, R. A., III; Sharp, N. Y.; Wood, D. A.	Accounting Horizons	Q3
20	2011	Nonaudit services provided by incumbent auditors and earnings management: Evidence of auditor independence from an EU country	Carmona Ibanez, P.; Momparler Pechuan, A.	Revista Espanola de Financiacion y Contabilidad-Spanish Journal of Finance and Accounting	Q4
21	2011	Recent Auditor Downgrade Activity and Changes in Clients' Discretionary Accruals	Carver, B. T.; Hollingsworth, C. W.; Stanley, J. D.	Auditing-A Journal of Practice & Theory	Q1
22	2011	Firm versus Partner Measures of Auditor Industry Expertise and Effects on Auditor Quality	Chi, H.; Chin, C.	Auditing-A Journal of Practice & Theory	Q1
23	2011	Litigation Risk and Abnormal Accruals	Boone, J. P.; Khurana, I. K.; Raman, K. K.	Auditing-A Journal of Practice & Theory	Q1
24	2010	Do Abnormally High Audit Fees Impair Audit Quality?	Choi, J.; Kim, J.; Zang, Y.	Auditing-A Journal of Practice & Theory	Q2

Tabela 4: Artigos incluídos na revisão sistemática entre 2009 a 1998

ID	Ano	Título	Autores	Journal	Quartil
25	2009	Accounting Discretion, Horizon Problem, and CEO Retirement Benefits	Kalyta, P.	Accounting Review	Q1
26	2008	Auditor Industry Specialization, Auditor Changes, and Accounting Restatements	Romanus, R. N.; Maher, J. J.; Fleming, D. M.	Accounting Horizons	Q3
27	2008	Audit partner tenure, audit firm tenure, and discretionary accruals: Does long auditor tenure impair earnings quality?	Chen, C.; Lin, C.; Lin, Y.	Contemporary Accounting Research	Q1
28	2008	Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical analysis	Jones, K. L.; Krishnan, G.I V.; Melendrez, K. D.	Contemporary Accounting Research	Q1
29	2006	Does hiring a new CFO change things? An investigation of changes in discretionary accruals	Geiger, M. A.; North, D. S.	Accounting Review	Q1
30	2005	Board monitoring and earnings management: Do outside directors influence abnormal accruals?	Peasnell, K. V.; Pope, R. F.; Young, S.	Journal of Business Finance & Accounting	Q2
31	2005	Do managers credibly use accruals to signal private information? Evidence from the pricing of discretionary accruals around stock splits	Louis, H; Robinson, D	Journal of Accounting & Economics	Q1
32	2005	Earnings management under price regulation: Empirical evidence from the Spanish electricity industry	Gill-De-Albornoz, B; Illueca, M	Energy Economics	Q1
33	2003	Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation?	Myers, J. N.; Myers, L. A.; Omer, T. C.	Accounting Review	Q1
34	2003	Earnings manipulation in failing firms	Rosner, R. L.	Contemporary Accounting Research	Q1
35	2002	Accruals management, investor sophistication, and equity valuation: Evidence from 10-Q filings	Balsam, S.; Bartov, E.; Marquardt, C.	Journal of Accounting Research	Q1
36	1998	Earnings management and the underperformance of seasoned equity offerings	Teoh, S.H.; Welch, I.; Wong, T. J.	Journal of Financial Economics	Q1

Após a realização da revisão dos artigos selecionados, estes foram alvo de avaliação da sua qualidade, tendo em conta o seu contributo para a concretização do objetivo e, em particular, para a resposta às quatro questões colocadas, apresentadas na Tabela 5.

Tabela 5: Critério de qualidade para avaliação dos artigos

Modelos baseados em <i>accruals</i>	Q0.1: Aborda os diferentes conceitos de manipulação de resultados?
	Q0.2: Descreve o método empírico baseado em <i>accruals</i> ?
	Q0.3: Compara o método baseados em <i>accruals</i> com outros métodos empíricos?
Tipologias de modelos	Q1: Faz a distinção entre <i>accruals</i> discricionários e <i>accruals</i> não discricionários?
	Q2: Descreve o(s) modelo(s) baseado(s) em <i>accruals</i> discricionários?
	Q3: Compara modelos baseados em <i>accruals</i> ?
	Q4: Descreve a aplicação dos modelos a demonstrações financeiras?
	Q5: Descreve o fim para que é aplicado o modelo?
Contexto de aplicação	Q6: O modelo é aplicado a uma amostra contextualizada?
	Q7: Justifica a escolha do modelo com base no tipo de amostra?
	Q8: A metodologia está devidamente apresentada?
	Q9: Justifica a eficácia do modelo no contexto da aplicação?
Avaliação dos resultados	Q10: Realiza técnicas de robustez e sensibilidade
	Q11: Descreve as limitações do estudo?
	Q12: Descreve os contributos do estudo?

2.2.1 Análise de rácios e tendências

A análise dos rácios e da sua evolução por diversos períodos constitui um método tradicional, sendo um dos métodos mais frequentes na análise da existência de práticas de MR. Os rácios financeiros, que constituem uma ferramenta de avaliação das demonstrações financeiras, estabelecem uma série de relações entre o balanço e a demonstração de resultados nas diferentes rubricas (Brigham & Houston, 2003).

Numa situação normal, é expectável que os indicadores espelhem um carácter de regularidade, ao longo dos diferentes períodos de relato. No entanto, caso este carácter regular seja descontinuado, é de esperar que esta irregularidade seja resultado de uma situação económica anómala, mas racional e justificável (Moreira, 2008). Na ausência deste motivo, as divergências no padrão de estabilidade devem ser consideradas potenciais indícios de manipulação (Cunha, 2013).

Penman (2012) propõe diferentes testes para deteção de MR, baseados na deteção de manipulação de vendas, manipulação de gastos e itens não recorrentes. No entanto, apesar da sua simplicidade, e da sua eficiência quando aplicada a empresas individualmente, esta metodologia apresenta severas limitações quando aplicada a empresas de maiores dimensões.

2.2.2. Observação e análise da distribuição dos resultados

Entre os principais incentivos à MR surge a tentativa de evitar relatos de resultados negativos, uma vez que a sua ocorrência pode acarretar para os gestores, penalizações por parte dos investidores/financiadores. Portanto, é expectável que os gestores, perante uma situação de prejuízo, se sintam impulsionados a transpor o resultado para uma situação de lucro ou mesmo neutra, através de práticas de MR (DeGeorge *et al.*, 1999).

Burgstahler e Dichev (1997), através da análise gráfica do histograma dos resultados líquidos, concluíram a existência de um pequeno número de empresas com pequenos prejuízos, ao que se contrapõe com uma frequência anormal elevada de empresas que relatam pequenos resultados positivos. A observada descontinuidade em torno de zero dos resultados, e a consequente quebra da normalidade da distribuição dos resultados que seria expectável dada a dimensão da amostra, sugerem a existência MR no relato financeiro.

Tal como no método da análise de rácios, a simplicidade surge como a grande vantagem deste método quando utilizado na deteção de MR. Quanto às limitações deste método, a pressuposição de que as irregularidades são apenas consequência da prática de gestão de resultados, e a ausência da identificação da forma utilizada para a MR, surgem como as principais questões levantadas pelos investigadores (Moreira, 2008).

2.2.3 Análise de *accruals*

Os *accruals* surgem quando existe uma discrepância entre o momento do reconhecimento do fluxo de caixa e o momento do resultado (Ronen & Yari, 2007), sendo definidos como a parte das receitas e despesas que ainda não se traduziu em recebimentos nem em pagamentos (Callao *et al.*, 2010).

Uma vez que, o resultado operacional do período de uma determinada empresa materializa-se em fluxo de caixa operacional (já concretizado o recebimento ou pagamento) e em *accruals* (ainda não configurado em recebimento ou pagamento), o cálculo dos *accruals* pode advir da diferença entre o resultado e os fluxos de caixa (Larcker *et al.*, 2004):

$$RO_{it} = CFO_{it} + ACC_{it} \quad (1)$$

Onde RO_{it} expressa o resultado operacional da empresa i no período t ; CFO_{it} representa o fluxo de caixa operacional da empresa i no período t ; ACC_{it} reflete os *accruals* da empresa i no período t .

Num mercado onde não existisse assimetria de informação, todos os *accruals* seriam resultado do normal funcionamento da empresa. No entanto, como tal pressuposto não se verifica no contexto económico atual, é necessário o cálculo de *accruals*, que pelo seu elevando nível de discricionariedade, pode levar à distorção da informação financeira.

Assim, os *accruals* totais correspondem à variação do fundo de maneio, incluindo as amortizações do período (Healy, 1985; Jones, 1991; Dechow *et al.* 1995). Em termos algébricos tem-se:

$$TA_t = \frac{\Delta CA_t + \Delta CL_t - \Delta Cash_t + \Delta STD_t - DEP_t}{A_{t-1}} \quad (2)$$

Onde TA_t : *accruals* totais no período t ; ΔCA_t : variação dos ativos correntes no período t face ao período $t - 1$; ΔCL_t : variação dos passivos correntes no período t face ao período $t - 1$; $\Delta Cash_t$: variação em caixa e equivalentes no período t face ao período $t - 1$; ΔSTD_t : variação da dívida, incluindo passivo corrente, no período t face ao período $t - 1$; DEP_t : total das depreciações e amortizações no período t ; A_{t-1} : ativo total no período $t - 1$.

Neste sentido, os *accruals* totais de uma empresa podem ser desagregados em *accruals* não discricionários (correspondem à componente explicada pelo normal funcionamento da empresa) e em *accruals* discricionários (correspondem à componente que não resulta da atividade real da empresa, sendo por isso interpretados como possíveis manifestações de atos de MR):

$$TA_{it} = AND_{it} + AD_{it} \quad (3)$$

onde TA_{it} expressa o total de *accruals* da empresa i no período t , e AND_{it} e AD_{it} expressam os *accruals* não discricionários e os *accruals* discricionários da empresa i no período t , respetivamente. Desta forma, quando se estabelece a relação entre *accruals* e MR, implicitamente está-se a referir aos AD.

Uma vez que as componentes não discricionárias e discricionárias dos *accruals* não são diretamente obtidos a partir dos documentos contabilísticos divulgados pela empresa (Dias, 2015),

Jones (1991) propôs um modelo que permite estimar a componente discricionária, e que ainda é considerado um modelo basilar, já que permite a generalização dos resultados, que modelos anteriores criados por Healey (1985), DeAngelo (1986) e Dechow e Sloan (1991) não permitiam. Adicionalmente, tratando-se da principal referência nesta temática da MR, o modelo de Jones foi sujeito a sucessivas modificações e tentativas de melhorias, das quais se destacam as introduzidas por Dechow *et al.* (1995), Larcker e Richardson (2004) e Kothari *et al.* (2005).

Por outro lado, como limitação deste modelo, surge a possibilidade de contaminação dos resultados, uma vez que os dados que servem de base à estimação do modelo podem ter sido alvo, por si só, de manipulação, o que resultaria em conclusões também distorcidas (McNichols, 2001), como também na dificuldade de obtenção de séries temporais longas para cada empresa (Jackson, 2018).

Assim, atendendo às inúmeras vantagens dos métodos baseados em *accruals* e acautelando as suas limitações, este estudo incide neste tipo de método. Importa referir que as suas limitações não são por si só impeditivas do sucesso deste método, e que não existe nenhum método isento de desvantagens, neste estudo optou-se por utilizar o método baseado em *accruals*, o mais difundido na literatura da especialidade.

2.3. Modelos baseados *accruals*

Na literatura empírica, quando se procura mensurar a prática de MR, muitos são os estudos que se fazem recorrer dos métodos baseados nas variações do capital circulante, também denominados de modelos baseados em *accruals* (*MBA*), sendo por isso um dos métodos dominantes nesta matéria de investigação.

2.3.1. Tipologias de modelos

Accruals trata-se de um método que tem como base a decomposição dos *accruals* totais, diretamente mensurados a partir das demonstrações financeiras, nas suas componentes discricionárias e não discricionárias. Desta forma, é necessário que os modelos criados procurem explicar e quantificar os *accruals* não discricionários, e através da diferença entre o TA e os AND possam ser mensurados os AD, ou seja aqueles que são passíveis de MR (Moreira, 2008).

McNichols (2000), no seu estudo sobre MR utilizando o método baseado em *accruals* identifica dois tipos de abordagens para a deteção da mesma. A primeira é através de modelos baseados em *accruals* agregados, que procuram estimar a componente discricionária dos *accruals* vistos de uma forma agregada (e.g., Jones, 1991; Dechow *et al.*, 1995). A segunda centra-se em modelos baseados em *accruals* específicos onde a sua abordagem consiste em identificar ou fazer uma modelização de

um modelo específico de *accruals*, tendo como objetivo analisar o comportamento das variáveis independentes que influenciam a prática de MR (e.g., McNichols & Wilson, 1988; Beneish, 1997).

Os primeiros modelos que procuram identificar este tipo de manipulação foram Healy (1985) e DeAngelo (1986), onde procuraram estimar os AND através da variação dos TA. Estes dois modelos usam a mesma *proxy*, isto é, utilizam os *accruals* totais como *proxy* da estimação dos AND, sendo que têm como pressuposto que estes são constantes ao longo do tempo, pois só assim é possível captar a variação dos AND. Healy (1985) sugere que a estimação dos AND de um dado ano seja obtida através da média aritmética simples dos TA registados num dado período de estimação, refletindo-se na seguinte expressão:

$$AND_t = \frac{\sum_{t=1}^T TA_t}{T} \quad (4)$$

Onde AND_t representa os *accruals* não discricionários, que são obtidos através da média dos TA (em que TA_t representa os *accruals* totais, relativos a T períodos). Desta forma, o valor de AD_t será dado pela diferença entre AND_t e TA_t .

Uma metodologia alternativa foi proposta por DeAngelo (1986), em que o autor considera que os AND no período t em análise deverão ser iguais aos TA do período anterior ($t - 1$). Neste sentido, se não existisse MR, a diferença entre os AND e os TA seria igual a zero. Desta forma este modelo representa-se algebricamente através da seguinte expressão:

$$AND_t = TA_{t-1} \quad (5)$$

Onde TA_{t-1} corresponde aos *accruals* totais do período anterior. A variável AND_t já se encontram definida anteriormente.

Na tentativa de ultrapassar as limitações apontadas aos modelos anteriores, Jones (1991) propõe um modelo que procura captar o efeito das circunstâncias económicas das empresas, rompendo, assim, com o pressuposto de modelos anteriores de que os AND seriam constantes ao longo do tempo. De facto, a autora sugere que os *accruals* são determinados em função da variação do volume de negócios da empresa e do seu ativo fixo tangível. Em concreto, o seu modelo é sintetizado pela seguinte equação:

Original Jones Model – Jones (1991)

$$TA_{i,t} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \alpha_2 \Delta REV_{i,t} + \alpha_3 PPE_{i,t} + \varepsilon_{i,t} \quad (6)$$

Onde $TA_{i,t}$ representa os *accruals* totais da empresa i no período t ; $\Delta REV_{i,t}$ simboliza a variação das vendas da empresa i no período t face ao período $t - 1$; $PPE_{i,t}$ corresponde ao ativo fixo tangível (*property, plant and equipment*) da empresa i no período t ; e $A_{i,t-1}$ refere-se ao valor total do ativo (*total assets*) da empresa i no período $t - 1$. $\varepsilon_{i,t}$ representa o erro aleatório, ou seja, tudo aquilo que não é explicado pelo modelo, da empresa i no período t .

Este modelo, ao integrar variáveis como as vendas e os ativos fixos tangíveis, tende a captar, pelo menos parcialmente, a especificidade de cada empresa. Por outro lado, ao considerar como variáveis independentes as vendas da empresa e o seu ativo fixo tangível, Jones (1991) parte do pressuposto de que a MR não é realizada com recurso a estes. No entanto, este modelo não está isento de limitações. Uma das principais lacunas apresentadas reside na incapacidade do modelo para capturar o impacto da manipulação nas vendas, por este assumir que estas estão associadas a *accruals* não discricionários (Dechow *et al.*, 1995).

Assim, Dechow *et al.* (1995), com base na limitação do modelo anterior, sugerem uma versão modificada do modelo original de Jones, usualmente denominado por *Modified Jones Model*, com o objetivo de eliminar a fonte de erro existente, através da seguinte expressão:

Modified Jones Model – Dechow *et al.* (1995)

$$TA_{i,t} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \alpha_2(\Delta REV_{i,t} - \Delta AR_{i,t}) + \alpha_3 PPE_{i,t} + \varepsilon_{i,t} \quad (7)$$

Onde $\Delta AR_{i,t}$ representa a variação das contas a receber da empresa i no período t face ao período $t - 1$. Todas as restantes variáveis já se encontram definidas anteriormente.

Neste modelo, os autores utilizam como ponto de partida a equação (6), ajustando o modelo proposto por Jones, retirando o efeito do crescimento das vendas a crédito, considerando-as discricionárias. Neste sentido, Dechow *et al.* (1995) partem do pressuposto que a MR numa situação de vendas a crédito é mais acessível do que em vendas a pronto pagamento. No entanto, este modelo não soluciona o problema para o qual foi criado, uma vez que a inclusão das vendas realizadas a crédito na equação, leva a que estas sejam consideradas discricionárias, o que poderá revelar-se inadequado em certas circunstâncias. Para além de Dechow *et al.* (1995), muitos foram os autores que criaram novas variações ao modelo *Modified Jones Model*.

O primeiro exemplo é o modelo proposto por Kasznick (1999) que faz acrescer mais uma variável explicativa ao *Modified Jones Model* (Dechow *et al.*, 1995), os *cash-flows* operacionais. Para o autor, a inclusão desta variável permite a mitigação de eventuais erros de medição dos *accruals*, uma vez que tem em consideração a *performance* operacional da empresa. O modelo faz-se representar pela seguinte equação:

Modified Jones Model with Cash-flow – Kasznick (1999)

$$TA_{i,t} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \alpha_2(\Delta REV_{i,t} - \Delta AR_{i,t}) + \alpha_3 PPE_{i,t} + \alpha_4 \Delta CFO_{i,t} + \varepsilon_{i,t} \quad (8)$$

Onde $\Delta CFO_{i,t}$ corresponde à variação dos fluxos de caixa operacionais da empresa i no período t face ao período $t - 1$. Todas as restantes variáveis já se encontram definidas anteriormente.

Larcker e Richardson (2004), por sua vez, para além do fluxo de caixa operacional já incluído no modelo anterior, acrescentam ainda como variável explicativa o rácio *book-to-mark* (BM). Esta variável

tem como objetivo captar o efeito das potenciais oportunidades de crescimento sobre os *accruals*. O modelo traduz-se na seguinte expressão:

Modified Jones Model with book-to-market ratio and cash flows – Larcker e Richardson (2004)

$$TA_{i,t} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \alpha_2(\Delta REV_{i,t} - \Delta AR_{i,t}) + \alpha_3 PPE_{i,t} + \alpha_4 BM_{i,t} + \alpha_5 CFO_{i,t} + \varepsilon_{i,t} \quad (9)$$

Onde $BM_{i,t}$ representa o rácio *book-to-market*³ da empresa i no período t . Todas as restantes variáveis já se encontram definidas anteriormente.

Larcker e Richardson (2004) apresentam evidência empírica de que o seu modelo tem um maior poder explicativo que o modelo de Jones (1991) modificado. Contudo, os autores continuam a assumir que todas as vendas a crédito são objeto de manipulação (Jones *et al.*, 2008).

Numa outra perspetiva, Kothari *et al.* (2005) referem que a estimação dos *accruals* deve ter em consideração a performance da empresa, pelo que acrescentam ao modelo modificado de Jones (Dechow *et al.*, 1995) a rentabilidade líquida do ativo (ROA). Desta forma, os autores introduzem o efeito rentabilidade do ativo, permitindo assim a sua comparabilidade entre empresas do mesmo setor. Deste modo, obtém-se a seguinte expressão:

Modified Jones Model with ROA – Kothari *et al.* (2005)

$$TA_{i,t} = \alpha_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \alpha_2(\Delta REV_{i,t} - \Delta AR_{i,t}) + \alpha_3 PPE_{i,t} + \alpha_4 ROA_{i,t} + \varepsilon_{i,t} \quad (10)$$

Onde $ROA_{i,t}$ corresponde à rentabilidade líquida do ativo⁴ (*return on assets*) da empresa i no período t . Todas as restantes variáveis já se encontram definidas anteriormente.

Para além dos modelos baseados em *accruals* discricionários, e segundo Jones *et al.* (2008), é importante incluir no estudo da MR a avaliação da qualidade dos *accruals*. Os autores defendem que modelos como o de Beneish (1997) permitem o estabelecimento de uma relação sistemática entre a probabilidade de manipulação e os dados das demonstrações financeiras selecionadas. Neste sentido, foram incluídos três modelos de modo a avaliar a qualidade dos *accruals*: Beneish (1997), Dechow e Dichev (2002) e McNichols (2002).

Beneish (1997), sugere um modelo que calcula um score, o M-Score, que resulta da combinação linear de oito variáveis financeiras que, analisadas isoladamente, não têm grande significado, mas que, em conjunto, procuram analisar que empresas provavelmente manipulam os seus resultados (Fernandes & Laureano, 2019).

³ Este rácio calcula-se através do quociente entre a capitalização bolsista e o capital próprio. A capitalização bolsista, por sua vez, é calculada através do produto do número de ações da empresa pela sua cotação em bolsa.

⁴ Calculado através do quociente entre o resultado líquido do período sobre o ativo total.

As oito variáveis que contribuem para este modelo encontram-se descritas na Tabela 6. O M-Score é obtido através da seguinte expressão:

M-Score – Beneish (1997)

$$M - SCORE_{i,t} = -4.840 + 0.920DSRI_{it} + 0.528GMI_{it} + 0.404AQI_{it} + 0.892SGI_{it} + 0.115DEPI_{it} - 0.172SGAI_{it} + 4.679TATA_{it} - 0.327LVGI_i \quad (11)$$

Tabela 6: Variáveis constituintes do M-Score de Beneish

Variável		Expressão de cálculo
DSRI	<i>Days Sales Receivable Index</i>	$\frac{AR_t/REV_t}{AR_{t-1}/REV_{t-1}}$
GMI	<i>Gross Margin Index</i>	$\frac{REV_{t-1}/REV_{t-1}}{REV_t - COGS_t/REV_t}$
AQI	<i>Asset Quality Index</i>	$\frac{1 - (Current\ assets_t + PPE_t)/AT_t}{1 - (Current\ assets_{t-1} + PPE_{t-1})/AT_{t-1}}$
SGI	<i>Sales Growth Index</i>	$\frac{REV_t}{REV_{t-1}}$
DEPI	<i>Depreciation Index</i>	$\frac{Depreciation_{t-1}/(Depreciation_{t-1} + PPE_{t-1})}{Depreciation_t/(Depreciation_t + PPE_t)}$
SGAI	<i>Sales, General, & Administrative Expenses Index</i>	$\frac{SGA_t/REV_t}{SGA_{t-1}/REV_{t-1}}$
TATA	<i>Total accruals to Total assets</i>	$(\Delta Current\ assets_t - \Delta Cash_t - (\Delta Current\ liabilities_t - \Delta Current\ maturities\ of\ long - term\ debt_t - \Delta Income\ tax\ payable_t) - Depreciation_t) / AT_t$
LVGI	<i>Leverage Index</i>	$\frac{(Long - term\ debt_t + Current\ liabilities_t)/AT_t}{(Long - term\ debt_{t-1} + Current\ liabilities_{t-1})/AT_{t-1}}$

Fonte: Beneish (1999:27)

Notas: COGS corresponde aos custos dos bens vendidos (*cost of goods sold*); PPE refere-se ao ativo fixo tangível (*property, plant and equipment*); SGA corresponde a *Selling, General & Administrative Expense*.

As variáveis deste modelo têm como finalidade capturar as distorções dos dados refletidos nas demonstrações financeiras, tendo sido agrupadas por Beneish (1999) e Beneish *et al.* (2013) em dois grupos distintos: indicadores de práticas de contabilidade agressiva (DSRI, DEPI e TATA) e indicadores de fraude (GMI, AQI, SGI, SGAI e LVGI).

O valor do M-score é obtido tendo em conta os valores determinados para cada uma das 8 variáveis que compõem a equação (11) (Beneish, 1999). Um M-score superior a -2,22 indica que as empresas em estudo apresentam uma elevada probabilidade de manipularem as suas demonstrações financeiras pelo que se o oposto se verificar indica que a empresa não é manipuladora (Beneish & Nichols, 2005).

Uma nova abordagem, proposta por Dechow e Dichev (2002), tem como pressuposto que os *accruals* ajustam o reconhecimento dos fluxos de caixa no tempo. Desta forma, e à semelhança dos modelos anteriores, as autoras mantêm a convicção que os AD são obtidos pela diferença entre os AT e os AND estimados. No entanto, e em oposição às abordagens anteriormente descritas, Dechow e Dichev (2002) consideram como variável independente a variação dos *accruals* de capital circulante em vez dos *accruals* totais, e como variáveis explicativas dos AND os fluxos de caixa da empresa. Desta forma este modelo representa-se algebricamente através da seguinte expressão:

Modelo de Dechow e Dichev (2002):

$$\Delta WC_{it} = \alpha_0 + \alpha_1 CFO_{it-1} + \alpha_2 CFO_{it} + \alpha_3 CFO_{it+1} + \varepsilon_{i,t} \quad (12)$$

Onde ΔWC_{it} corresponde à variação do fundo de maneo⁵ (*working capital*) da empresa *i* no período *t* face ao período *t - 1*; e CFO_{it-1} ; CFO_{it} e CFO_{it+1} expressam os fluxos de caixa da empresa *i* no período de tempo *t - 1*, *t* e *t + 1*, respetivamente. Todas as restantes variáveis já se encontram definidas anteriormente.

No entanto, a principal crítica apontada ao modelo de Dechow e Dichev (2002) sugere que as variáveis explicativas do modelo estão sujeitas a problemas de mensuração que podem dar origem a coeficientes enviesados.

Assim, McNichols (2002), e de forma a ultrapassar a limitação do modelo anterior o problema, propõe um modelo onde procura relacionar a estratégia de Jones (1991) com as variáveis explicativas propostas por Dechow e Dichev (2002), fazendo-se representar pela seguinte expressão:

Modelo de McNichols (2002):

$$\Delta WC_{it} = \alpha_0 + \alpha_1 CFO_{it-1} + \alpha_2 CFO_{it} + \alpha_3 CFO_{it+1} + \alpha_4 \Delta REV_{it+} + \alpha_5 PPE_{it} + \varepsilon_{i,t} \quad (13)$$

Todas as variáveis já se encontram definidas anteriormente.

O autor propõe como variáveis explicativas, a utilização não só dos fluxos operacionais da empresa (do ano anterior, do presente e do ano seguinte: *t - 1*, *t* e *t + 1* respetivamente), como também da variação das vendas e do ativo fixo tangível (*property, plant and equipment*), uma vez que considera que estas variáveis irão aumentar o poder explicativo do modelo.

A Tabela 7 explicita as tipologias de modelos analisados dos artigos alvo da RSL assinalando se descreve e/ou aplica os 10 modelos anteriormente apresentados. Entre os 36 artigos incluídos nesta revisão sistemática, o artigo de Jones *et al.* (2008) destaca-se por descrever e aplicar 7 dos 10 MBA, sendo este o artigo que expõe o maior número de modelos. Em oposição, o artigo de Romanus *et al.* (2008) apenas aplica um modelo, sem especificar que modelo é utilizado, e sem apresentar uma breve descrição do mesmo. Do número total de artigos, existe uma tendência clara para o número de

⁵ Calculado através da diferença entre o ativo corrente e o passivo corrente.

modelos descritos ser o mesmo de modelos aplicados ($n = 25$) (Lin & Liao, 2018; Geiger & North, 2006).

Tabela 7: Tipologias de modelos utilizadas nos artigos da RSL

Artigo	HEALY		DEANGELO		OJONES		MJONES		MJONES2		MJONES3		MJONES4		BENEISH		DECHOW		MCNICHOLS		Total ¹	
	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A
1							✓	✓					✓	✓			✓		✓	✓	4	3
2							✓		✓	✓			✓	✓							3	2
3					✓	✓											✓	✓			2	2
4					✓	✓							✓								2	1
5							✓	✓					✓	✓							2	2
6							✓	✓					✓	✓							2	2
7													✓	✓							1	1
8													✓	✓							1	1
9													✓	✓			✓	✓			2	2
10					✓			✓			✓	✓	✓	✓			✓	✓			3	5
11							✓	✓					✓	✓			✓	✓	✓		4	3
12													✓	✓			✓	✓			2	2
13							✓	✓					✓								2	1
14							✓	✓					✓	✓							2	2
15							✓		✓	✓											2	1
16					✓	✓	✓	✓													2	2
17					✓								✓	✓			✓	✓	✓	✓	4	3
18					✓	✓	✓	✓													2	2
19							✓	✓							✓	✓	✓	✓			3	3
20					✓		✓						✓	✓							3	1
21					✓	✓	✓	✓					✓	✓							3	3
22							✓						✓	✓							2	1
23					✓	✓							✓	✓							2	2
24							✓	✓					✓	✓							2	2
25							✓	✓													1	1
26																					0*	0*
27							✓	✓					✓	✓							2	2
28					✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	7	7
29					✓	✓							✓	✓							2	2
30					✓		✓	✓													2	1
31							✓	✓					✓	✓							2	2
32					✓	✓			✓	✓											2	2
33					✓	✓															1	1
34	✓	✓			✓	✓	✓	✓													3	3
35					✓	✓															1	1
36					✓	✓	✓	✓													2	2
Total²	1	1	0	0	16	14	22	19	3	3	2	2	23	21	2	2	9	8	4	3		

Notas: D: descreve; A: aplica; HEALY: Modelo defendido por Healy em 1985; DEANGELO: Modelo defendido por DeAngelo em 1986; OJONES: Modelo defendido por Jones em 1991; MJONES: Modelo defendido por Dechow *et al.* em 1995; MJONES2: Modelo defendido por Kasznick em 1999; MJONES3: Modelo defendido por Larcker e Richardson em 2004; MJONES4: Modelo defendido por Kothari *et al.* em 2005; BENEISH: Modelo defendido por Beneish em 1997; DECHOW: Modelo defendido por Dechow e Dichev em 2002; MCNICHOLS: Modelo defendido por McNichols em 2002. Total¹: Número total de modelos descritos e/ou aplicados em cada artigo. Total²: Número total de artigos que aplicam cada modelo.

Entre os modelos avaliados, o modelo de DeAngelo (1986), e o modelo de Healy (1985) são os modelos que menos foram referidos, sendo que o primeiro nunca fora mencionado, e o segundo apenas foi mencionado uma vez (Rosner, 2003). Os mais mencionados foram o de Jones (1991),

Dechow (1995), Kothari (2005), sendo o artigo de Carver *et al.* (2011), um exemplo disso uma vez que descreveu e aplicou os três modelos.

2.3.3. Contextos de aplicação

Os artigos incluídos na RSL, na sua maioria, têm como âmbito de estudo a auditoria (Salehi, 2018; Myers *et al.*, 2003); e o *reporting* (Jha, 2019; Louis & Robinson, 2005), sendo a fiscalidade um tema abordado apenas por um dos artigos estudados (Wong *et al.*, 2015). Isto reflete um crescente interesse sobre a qualidade dos *reports* emitidos pelas empresas, e de que modo a auditoria tem impacto na qualidade nos *accruals*. A amostra destes 36 artigos é composta em média por 1559 empresas, tendo sido alvo de estudo por um período médio de 13 anos (De Fuentes *et al.*, 2015). Ambas as variáveis são bastante heterogéneas, uma vez que o número de empresas que compõe a amostra dos artigos analisados varia entre 13 (Gill-De-Albornoz, & Illueca, 2005) e 28289 (Boone *et al.*, 2011), e o intervalo temporal de análise tem como extremos os 3 anos (Carver *et al.*, 2011; Balsam *et al.*, 2002) e os 30 anos (Simpson, 2013).

Esta amostra, caracteriza-se ainda, por serem na sua maioria empresas cotadas, e onde são incluídos mais que um setor de atividade, sendo que apenas 2 dos 36 o estudo é focado em apenas um setor (Beck, 2018; Gill-De-Albornoz & Illueca, 2005). No entanto, grande parte destes artigos, excluí prontamente as instituições financeiras da amostra, de forma a eliminar o enviesamento dos resultados que poderá existir devido aos efeitos da própria regulação deste setor. Como alvo de estudo, destacam-se as empresas sediadas nos Estados Unidos da América (Kalyta, 2009; Teoh *et al.*, 1998)), seguidas de Espanha (Carmona Ibanez & Momparler Pechuan, 2011) e Reino Unido (Peasnell *et al.*, 2005), fazendo com que a base de dados mais utilizada seja a COMPUSTAT, uma vez que se trata de uma base de dados Americana (Fisher *et al.*, 2019; Chang & Lin, 2018; Choi *et al.*, 2010).

Em relação à técnica de análise utilizada, todos os artigos utilizaram pelo menos o método de regressão linear (Chi & Chin, 2011; Chen *et al.*, 2008; Geiger & North, 2006), podendo esta ser estimada em OLS, ou em dados em painel. A este método, 8 dos 36 artigos incluíram também a regressão logística (Jha, 2019; Peasnell *et al.*, 2005), e 11 recorreram aos testes de hipótese (Qi, 2017; Kurt, 2018). Apenas 2 dos artigos estudados utilizaram tanto a regressão linear, a regressão logística e os testes de hipóteses como principal metodologia (Qi, *et al.*, 2017; Jones *et al.*, 2008). Desta forma, os 36 artigos analisados refletem uma tendência para que os *accruals* sejam integrados nas regressões como variáveis dependentes, o que corrobora com o facto da principal finalidade dos mesmos é a sua mensuração (Ravenda *et al.*, 2018; Habib & Bhuiyan, 2016). Em oposição, 6 dos 36 artigos acresce ainda o objetivo de avaliar a qualidade dos *accruals* (Chi & Chin, 2011) fortalecendo a ideia já defendida por Jones *et al.* (2008) anteriormente. A Tabela 8 resume o contexto de cada investigação abordada pelos artigos alvo da RSL.

Tabela 8: Contextualização da aplicação da MR dos artigos da RSL

ID	Âmbito	Nº de empresas	Período	Tipo de Empresas	Setor	País da amostra	Base de dados	Técnica de Análise	ACC como V.I.	Principal finalidade
1	Reporting	243	2005-2012	Não Cotadas	>1	Coreia do Sul	Data Guide5; KRX's KIND	Regressão Linear	Não	Qualidade e Mensuração
2	Bankruptcy	261	1995-2009	N.E.	>1	EUA	COMPUSTAT	Regressão Linear	Não	Mensuração
3	Reporting	10168	1990-2009	Não Cotadas	>1	EUA	Federal Securities Regulation; COMPUSTAT	Regressão Logística e regressão Linear	Não	Qualidade e Mensuração
4	Contabilidade Pública	232	2008-2013	Não Cotadas	1	EUA	United States Census Bureau; SDC Platinum	Regressão Linear e Teste de hipóteses	Não	Mensuração
5	Reporting	543	1986-2012	Cotadas	>1	EUA	CRSP IBES	Regressão Linear e Teste de hipóteses	Sim	Mensuração
6	Seasoned equity offerings	2977	1987-2012	Cotadas	>1	EUA	COMPUSTAT	Regressão Linear e Teste de hipóteses	Não	Mensuração
7	Auditoria	172	2009-2016	Cotadas	>1	Irão	N.E.	Regressão Logística e regressão Linear	Não	Mensuração
8	Seasoned equity offerings	1317	1983-2014	Cotadas	>1	EUA	COMPUSTAT	Regressão Linear e Teste de hipóteses	Não	Mensuração
9	Fraude	224	2003-2012	Não Cotadas	>1	Itália	ANBSC; AIDA	Regressão Linear e Teste de hipóteses	Não	Mensuração
10	N.E.	10369	1988-2007	Cotadas	>1	EUA	COMPUSTAT; CRSP	Regressão Linear e Teste de hipóteses	Sim	Mensuração
11	Auditoria	1066*	1999-2010	N.E.	>1	China	CSMAR	Regressão Logística, Regressão Linear e Teste de hipóteses	Não	Mensuração
12	Reporting	7040 3338	2004-2010	N.E.	>1	EUA	COMPUSTAT	Regressão Linear e Teste de hipóteses	Não	Mensuração
13	Auditoria	193* 326*	1995-2007	Não Cotadas	N.E.	Espanha	IBSAS	Regressão Logística e regressão Linear	Não	Qualidade e Mensuração
14	Reporting	644	1984-2005	Cotadas	>1	EUA	COMPUSTAT; CRSP	Regressão Linear e Teste de hipóteses	Sim	Mensuração
15	Fiscalidade	438	2001-2006	Cotadas	>1	China	Wind Financial; CSMAR	Regressão Linear	Não	Mensuração
16	Seasoned equity offerings	5573*	1976-2005	Cotadas	>1	EUA	COMPUSTAT; CRSP	Regressão Logística e regressão Linear	Não	Mensuração
17	Auditoria	257*	2005-2007	Cotadas	>1	Finlândia e Suécia	Thomson Reuters Worldscope	Regressão Linear e Teste de hipóteses	Não	Mensuração
18	Bankruptcy	343*	1994-2004	Cotadas	>1	Reino Unido	Datastream	Regressão Logística e regressão Linear	Não	Mensuração
19	Fraude	4085*	1995-2008	N.E.	>1	EUA	COMPUSTAT; CRSP	Regressão Linear	Sim	Qualidade e Mensuração

ID	Âmbito	Nº de empresas	Período	Tipo de Empresas	Setor	País da amostra	Base de dados	Técnica de Análise	ACC como V.I.	Principal finalidade
20	Auditoria	69*	2005-2009	Cotadas	>1	Espanha	SABI	Regressão Linear	Não	Mensuração
21	Auditoria	952	2003-2005	N.E.	>1	EUA	COMPUSTAT	Regressão Linear	Não	Mensuração
22	Auditoria	370* 403*	1983-2004	Cotadas	>1	Taiwan	Taiwan Economic Journal	Regressão Linear	Não	Qualidade e Mensuração
23	Auditoria	67 28289	1989-2007	N.E.	>1	EUA	COMPUSTAT; CRSP	Regressão Linear	Não Sim	Mensuração
24	Auditoria	1765*	2000-2003	N.E.	>1	EUA	COMPUSTAT	Regressão Linear	Não	Mensuração
25	Reporting	388	1997-2006	Não se aplica	>1	EUA	COMPUSTAT	Regressão Linear	Não	Mensuração
26	Auditoria	456	1998-2003	Cotadas	>1	EUA	COMPUSTAT	Regressão Linear	Sim	Qualidade e Mensuração
27	Auditoria	888	1990-2001	Cotadas	>1	Taiwan	Taiwan Economic Journal	Regressão Linear	Não	Mensuração
28	Fraude	118	1988-2001	N.E.	>1	EUA	COMPUSTAT	Regressão Logística, Regressão Linear e Teste de hipóteses	Sim	Mensuração
29	Reporting	712	1994-2000	Cotadas	>1	EUA	COMPUSTAT	Regressão Linear	Não	Mensuração
30	Reporting	424*	1993-1996	Cotadas	>1	Reino Unido	London Share Price; Corporate Register; Datastream	Regressão Logística e regressão Linear	Não	Mensuração
31	Reporting	175*	1990-2002	Cotadas	>1	EUA	COMPUSTAT; CRSP	Regressão Linear	Sim	Mensuração
32	Regulação	13	1991-2001	Cotadas	1	Espanha	N.E.	Regressão Linear e Teste de hipóteses	Não	Mensuração
33	Auditoria	3254*	1988-2000	N.E.	>1	EUA	COMPUSTAT	Regressão Linear	Não	Mensuração
34	Bankruptcy	200*	1985-1997	N.E.	>1	EUA	COMPUSTAT Compact Disclosure	Regressão Linear e Teste de hipóteses	Não	Mensuração
35	Reporting	366	1996-1998	Cotadas	N.E.	EUA	COMPUSTAT; IBES; EDGAR; Compact Disclosure; CRSP	Regressão Linear	Não	Mensuração
36	Seasoned equity offerings	1265	1976-1989	Cotadas	>1	EUA	COMPUSTAT; CRSP	Regressão Linear	Não	Mensuração

Notas: N.E: Não especificado; *Seasoned equity offerings* = aumento da capital para empresas públicas; *calculado através do quociente entre o número de observações ano-empresa sobre o período de estudo. CRSP: Center for Research in Security Prices; IBES: Institutional Brokers Estimate System; ANBSC: National Agency for the Management and Assignment of Seized and Confiscated Assets; AIDA: Italian Bureau Van Dijk database; CSMAR: China Security Market & Accounting Research; IBSAS: Iberian Balance Sheet Analysis System; SABI: Balance Sheet Analysis System for Iberian Companies; EDGAR: Electronic Data Gathering, Analysis, and Retrieval system.

2.3.2. Avaliação dos resultados

Com a intenção de avaliar os resultados obtidos nos modelos, a Tabela 9 sistematiza os 36 artigos da RSL. Foram alvo de análise os testes de sensibilidade/robustez realizados, as limitações dos estudos, e os seus contributos.

Quanto aos testes de sensibilidade foram observados cinco âmbitos diferentes em que estes foram aplicados, sendo que em todos aos artigos foram utilizados testes de robustez. Assim, houve artigos que apenas utilizaram um único âmbito de teste (Teoh *et al.* 1998; Chu-Hsuan & Hsiou-Wei William, 2018), enquanto que outros preferiram utilizar mais do que um (Kalyta, 2009), num máximo de três (Jones *et al.*, 2008; Wong, 2015).

Dos cinco âmbitos utilizados, os mais recorrentes são a alteração das variáveis independentes, onde por exemplo, foram adicionadas outras variáveis de controlo à regressão logística inicial (Atieh & Hussain, 2012) e a utilização de outros métodos de regressão, por exemplo, a regressão linear através do método de estimação dos mínimos quadrados ordinários (OLS) (Jha, 2019). Por outro lado, houve artigos que, de forma a comprovar a robustez do seu estudo, aplicaram a sua metodologia a uma amostra diferente, quer em termos de empresas e/ou em termos de período de estudo. Mais concretamente, Perotti & Windisch (2017), decidiram estudar a estabilidade dos resultados obtidos inicialmente, aplicando a sua regressão OLS à amostra inicial, agora repartida em dois períodos de 10 anos, e em quatro períodos de 5 anos. Noutra perspetiva, Ravenda *et al.* (2018) preferiram aplicar as suas técnicas de análise de dados a uma outra amostra, isto é, a um novo conjunto de empresas. De forma menos expressiva, surgem as alterações às variáveis dependentes, onde a equação inicial foi re-estimada tendo em conta quatro diferentes métricas de auditoria (Romanus *et al.*, 2008), e a utilização de outro modelo de *accruals*, neste caso foi medida a qualidade dos *accruals* segundo o modelo de Dechow e Dichev (Ahsan & Md. Borhan Uddin, 2016).

Quanto às limitações, muitos dos artigos não as especificaram, no entanto, quando o faziam, as mais referidas foram os dados, mais precisamente, o elevado número de não respostas (Carver, *et al.* 2011) e o período de estudo, onde Beck (2018) refere que o facto do seu estudo apenas ter recaído num período de apenas seis anos condicionou os seus resultados.

Por fim, os contributos foram categorizados em duas dimensões: nova dimensão de análise, isto é, quando o estudo aborda um novo tema ou perspetiva em relação há bibliografia já existentes (Carmona Ibanez & Momparler Pechuan, 2011) e aprofundamento do âmbito, onde a investigação procura explorar questões já abordadas anteriormente (Lin & Liao, 2018).

Tabela 9: Avaliação dos resultados dos artigos alvo da RSL

ID	Testes de sensibilidade/robustez	Limitações	Contributos
1	Outros métodos de regressão	N.E.	Nova dimensão de análise
2	Outros métodos de regressão	N.E.	Nova dimensão de análise
3	Outros métodos de regressão	N.E.	Nova dimensão de análise
4	Alteração das VI	Dados e período	Nova dimensão de análise
5	Outros métodos de regressão e alteração das VI	N.E.	Aprofundamento do âmbito
6	Alteração das VI	Dados	Nova dimensão de análise; Aprofundamento do âmbito
7	Outros métodos de regressão	N.E.	Nova dimensão de análise
8	Outros métodos de regressão	Dados	Aprofundamento do âmbito
9	Alteração da amostra (empresas)	Dados	Nova dimensão de análise
10	Outros métodos de regressão, alteração da amostra (empresas e período)	Dados	Aprofundamento do âmbito
11	Alteração da amostra (período); outro modelo de <i>accruals</i>	Dados	Aprofundamento do âmbito
12	Outro modelo de <i>accruals</i>	N.E.	Aprofundamento do âmbito
13	Alteração da amostra (empresas); alteração das VI	Dados e período	Aprofundamento do âmbito
14	Outros métodos de regressão	Dados	Aprofundamento do âmbito
15	Alteração da amostra (período); alteração da VD; alteração das VI	N.E.	Aprofundamento do âmbito
16	Alteração da amostra (período); alteração das VI	N.E.	Aprofundamento do âmbito
17	Alteração da VD; alteração das VI	Dados	Aprofundamento do âmbito
18	Alteração das VI	Dados	Aprofundamento do âmbito
19	Outros métodos de regressão; alteração das VI	Dados	Nova dimensão de análise
20	Alteração das VI	N.E.	Nova dimensão de análise
21	Outros métodos de regressão	Dados	Aprofundamento do âmbito
22	Alteração das VI	N.E.	Aprofundamento do âmbito
23	Alteração da VD; alteração das VI	N.E.	Nova dimensão de análise; Aprofundamento do âmbito
24	Alteração da amostra (período); alteração das VI	Dados	Aprofundamento do âmbito
25	Alteração das VI	Dados	Aprofundamento do âmbito
26	Alteração da VD; alteração das VI	Dados	Aprofundamento do âmbito
27	Alteração das VI; Outro modelo de <i>accruals</i>	Dados	Nova dimensão de análise; Aprofundamento do âmbito
28	Alteração das VI; Outro modelo de <i>accruals</i> ; Alteração da amostra (empresas e período)	N.E.	Aprofundamento do âmbito
29	Outro modelo de <i>accruals</i>	N.E.	Aprofundamento do âmbito
30	Alteração das VI	N.E.	Nova dimensão de análise
31	Alteração das VI	N.E.	Nova dimensão de análise; Aprofundamento do âmbito
32	Alteração das VI; Outro modelo de <i>accruals</i>	N.E.	Nova dimensão de análise; Aprofundamento do âmbito
33	Alteração da amostra (empresas e período)	N.E.	Nova dimensão de análise
34	Alteração das VI	N.E.	Aprofundamento do âmbito
35	Alteração das VI	N.E.	Nova dimensão de análise; Aprofundamento do âmbito
36	Outros métodos de regressão	N.E.	Aprofundamento do âmbito

Notas: VI: Variável independente; VD: variável dependente; N.E: não específica.

2.4 Avaliação dos artigos científicos relevantes

Uma vez que este estudo tem por base uma revisão sistemática, é importante facilitar o trabalho a investigadores e profissionais, pelo que se avalia a relevância dos artigos em cada um dos seguintes quatro aspetos dos métodos e modelos tal como apresentado, anteriormente, na Tabela 5. Para cada critério de qualidade é atribuído uma pontuação a cada pergunta de avaliação e para cada artigo: 0 se não responde à questão; 0,5 se responde parcialmente e 1 se responde totalmente. A Tabela 10 avalia os 36 artigos sujeitos ao estudo.

Tabela 10: Avaliação da qualidade dos artigos que integram a RSL

ID	Métodos baseados em <i>accruals</i>			Tipologias de modelos				Contexto de aplicação					Avaliação dos resultados			Total ¹
	Q0.1	Q0.2	Q0.3	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	
1	0	0	0	0	1	1	1	1	1	1	1	1	1	0	1	10
2	0,5	1	0,5	0	1	1	1	1	1	1	1	1	1	0	1	12
3	0,5	0	0	0	0,5	1	0	1	1	1	1	1	1	0	1	9
4	0,5	0,5	0	0	0,5	0	0	1	1	1	1	1	1	1	1	9,5
5	0,5	0,5	0	0	1	1	1	1	1	1	1	1	1	0	1	11
6	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
7	0	0	0	0	1	0	1	1	0,5	1	0,5	1	1	0	1	8
8	0	0	0	0	0,5	0	0	1	1	0,5	1	1	1	1	1	8
9	0	0	0	0	1	1	1	1	1	1	1	0,5	1	1	1	10,5
10	0	0,5	0	0	0,5	1	0	0,5	1	0,5	1	1	1	1	1	9
11	0	0,5	0	0	0,5	1	0,5	1	1	0	1	1	1	1	1	9,5
12	0	0,5	0	0	0,5	1	1	1	0,5	1	1	1	1	0	1	9,5
13	0	0,5	0	0	0,5	1	0,5	1	1	1	1	1	1	1	1	10,5
14	0	0	0	0	1	1	0,5	1	1	1	1	1	1	0	1	9,5
15	0	0	0	0	0,5	0,5	0,5	1	1	1	1	1	1	0	1	8,5
16	0	1	0	0,5	0,5	1	1	1	1	1	1	1	1	0	1	11
17	0	0	0	0	0,5	1	1	1	1	1	1	1	1	1	1	10,5
18	0,5	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11,5
19	0,5	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11,5
20	0	0,5	0	0	1	0,5	1	1	1	0,5	1	1	1	0	1	9,5
21	0	0,5	0	0	0,5	1	1	1	1	1	1	0,5	1	1	1	10,5
22	0	0,5	0	0	0,5	0,5	1	1	1	0,5	1	1	1	0	1	8,5
23	0	0	0	0	0,5	1	1	1	1	1	1	0,5	1	0	1	9
24	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
25	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	9
26	0	0	0	0	0	0	1	1	1	1	1	0,5	1	1	1	8,5
27	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
28	0	1	0,5	0,5	1	1	1	1	1	1	1	1	1	0	1	12
29	0	0	0	0	0,5	1	1	1	1	1	1	1	1	0	1	9,5
30	0	0	0	1	1	0,5	1	1	1	1	1	1	1	0	1	10,5
31	0	0	0	0	1	1	1	1	1	1	1	1	1	0	1	10
32	0	0	0	0,5	1	1	1	1	1	1	0,5	0,5	1	0	1	9,5
33	0	0	0	1	1	0	1	1	1	1	1	1	1	0	1	10
34	0	0	0	0	0	0,5	1	1	1	0,5	0,5	1	1	0	1	7,5
35	0	0	0	0	0,5	0	1	1	1	1	1	1	1	0	1	8,5
36	0	0	0	0,5	1	1	1	1	1	1	1	1	1	0	1	10,5
Total²	3	7,5	1	4	26,5	26,5	29	35,5	35	32,5	34,5	33,5	36	15	36	

Notas: Total¹: Pontuação total de cada artigo. Total²: Pontuação total de cada critério de qualidade.

Dos artigos incluídos nesta RSL, conclui-se que apenas dois dos artigos analisados obtiveram maior pontuação em termos de qualidade, 11,5/15, (Fisher *et al.* 2019; Jones *et al.* 2008), sendo que o critério de avaliação que apresenta maior qualidade é o Q5.

A dimensão que se encontra com melhor avaliação é o contexto de aplicação, onde todos os artigos abordam as diversas questões colocadas, no entanto, a avaliação dos resultados não fica aquém. Aliás, é nesta dimensão onde se encontram os dois critérios com melhor *score* (Q10 e Q12), onde são alvo de análise os testes de sensibilidade/robustez e os contributos do estudo. Em oposição, surge a primeira dimensão a ser avaliada, os métodos baseados em *accruals*, onde a comparação do MBA com outros métodos empíricos (Q0.3), surge com uma avaliação muito baixa. Neste sentido, conclui-se que poucos são os artigos que abordam outros métodos empíricos para a deteção de *accruals*, como também aqueles que refletem acerca dos vários conceitos existentes para definir manipulação de resultados (Q0.1).

3. Metodologia

Neste capítulo são explicitados os procedimentos metodológicos adotados ao desenvolvimento deste estudo. Essencialmente, são apresentados o posicionamento e a teoria de suporte da investigação, bem como o processo de recolha dos dados e as técnicas de análise dos dados.

3.1 Posicionamento da Investigação

A amplificação da investigação na área da contabilidade, em muito tem contribuído para a diversificação e desenvolvimento de abordagens e respetivos paradigmas (Major, 2009). Se até aos anos de 1960 o paradigma teórico predominante será o normativo, caracterizado pela maximização dos proveitos e o processamento de informação de uma forma perfeita, a partir de 1980 a pesquisa positivista ou empírica popularizou-se. Esta permite testar relações entre as diversas variáveis, expressas em hipóteses formuladas previamente (Major, 2017).

Tendo em conta os objetivos gerais e específicos do presente estudo, é coerente afirmar que a metodologia adotada segue um género de pesquisa quantitativa, que, por sua vez, se integra no paradigma positivista. Este estudo caracteriza-se por se tratar de um estudo exploratório, uma vez que ao abordar a MR em fundos europeus, visa obter o conhecimento de realidade pouco abordada e levantar hipóteses de entendimento dessa realidade; assentar num processo dedutivo, isto é, são formuladas hipóteses com base na literatura e validadas para a amostra selecionada, mais concretamente nas candidaturas a fundos; e utilizar dados considerados numéricos para validar hipóteses e modelos de estudo (Sousa & Batista, 2011).

3.2 Amostra e dados

A informação financeira deste estudo foi recolhida na Base de Dados *Amadeus All Companies*, que pertence à *Bureau van Dijk (BvD)*, onde é partilhada informação sobre as maiores 555.000 empresas públicas e privadas, a nível Europeu, sendo abrangidos 43 países. Assim, e através da BvD foi possível obter toda a informação financeira necessária ao estudo, como também proceder à sua caracterização. Esta informação é conjugada com a informação disponibilizada pelo IAPMEI, entidade responsável por parte dos incentivos financeiros atribuídos às empresas, relativa às candidaturas recebidas. Daqui resultam todos os dados, após garantida a sua qualidade e preparação, utilizados nas análises que permitem atingir os objetivos.

3.2.1 Processo de seleção da amostra

Uma vez que o presente estudo procura investigar a MR em fundos europeus, a amostra é constituída por candidaturas de empresas a fundos europeus entre os anos 2012 e 2019. Numa primeira instância, a recolha de dados foi realizada recorrendo ao IAPMEI, que disponibilizou os dados das empresas que

se candidataram a incentivos europeus neste período, o que originou uma amostra inicial de 28.658 candidaturas, às quais correspondem 19.087 empresas.

Posteriormente, e recorrendo à base de dados *Bureau van Dijk Amadeus*, foi recolhida informação financeira e relatórios de *corporate governance* de 16.672 empresas que correspondem a 22.686 candidaturas, para o período compreendido entre 2010 e 2020. Neste sentido, foram desconsideradas 5.972 candidaturas por ausência de informação e 264 por representarem uma duplicação de candidatura, ou seja, empresas que se candidataram ao mesmo projeto mais que uma vez. Assim, este estudo considera uma amostra de 22.422 candidatura, que se fazem corresponder a 14.122 empresas. Em suma, a amostra selecionada teve como triagem o procedimento apresentado na Tabela 11.

Tabela 11: Processo de seleção da amostra

Descrição	Nº candidaturas	Nº empresas
Base de dados IAPMEI	28.658	19.087
Após eliminação de candidaturas sem informação financeira suficiente na Amadeus	22.686	16.672
Após eliminação de candidaturas duplicadas	22.422	14.122

3.2.2 Caracterização da amostra

A amostra selecionada é composta por 22.422 candidaturas a fundos europeus entre 2012 e 2019. Em primeiro lugar, a Tabela 12 visa refletir a distribuição das candidaturas constituintes da amostra, por setor de atividade. Esta distribuição tem como base a Classificação Portuguesa de Atividades Económicas, Revisão 3 (CAE-Rev.3), sendo que esta representa o novo enquadramento das atividades económicas portuguesas, no sentido de a harmonizar com a Nomenclatura Estatística das Atividades Económicas da Comunidade Europeia (INE, 2007). Neste sentido, e para garantir representatividade dos diferentes setores, optou-se por agrupar os setores da secção C à F, e da H à S, designando-os por Indústria e Serviços, respetivamente.

Tabela 12: Distribuição de candidaturas por setor de atividade entre 2012 e 2019

Secção	Descrição	SIC Code	Nº Obs	%
A	Agricultura, produção animal, caça, floresta e pesca	01-03	298	1,33
B	Indústrias extrativas	05-09	153	0,68
C à F	Indústria	10-43	11.389	50,79
G	Comércio	45-47	4.183	18,66
H à S	Serviços	49-96	6.399	28,54

A secção com maior expressão neste estudo é o setor da Indústria, representando cerca de metade das candidaturas (50,79%), seguindo-se o setor dos serviços com um contributo de 28,54% para a amostra.

Noutro sentido, a Tabela 13 evidencia que as regiões onde se concentram um maior número de candidaturas são o Norte ($n = 11.127$) e o Centro ($n = 7.373$), o que contrasta com os números das regiões autónomas, a Madeira ($n = 13$) e Açores ($n = 1$), onde estes são pouco expressivos. Quanto à dimensão, constata-se que são as PME aquelas que assumem maior preponderância, com 97,61% ($n = 21.887$), opondo-se às “Não PME” que representam apenas 2,18% da amostra ($n = 489$). Não obstante, surge ainda uma quinta categoria, “Não aplicável”, com apenas 46 candidaturas, uma vez que neste estudo fazem parte da amostra candidaturas de associações e de empresas com forma legal indefinida.

Tabela 13: Distribuição de candidaturas por NUTS II e dimensão entre 2012 e 2019

NUTS II/ Dimensão	Não Aplicável	Não PME	Média	Pequena	Micro	Total
Alentejo	1	21	218	482	764	1.486
Algarve	0	1	40	170	418	629
Centro	29	137	1.416	2.744	3.047	7.373
Lisboa e Vale do Tejo	12	109	312	629	731	1.792
Norte	4	221	2.248	4.142	4.512	11.127
Região Autónoma da Madeira	0	0	3	3	7	13
Região Autónoma dos Açores	0	0	0	0	1	1
Total	46	489	4.237	8.170	9.480	22.422

3.3 Caracterização e mensuração dos dados a analisar

Este estudo procura verificar a existência de relação entre a manipulação de resultados, e a elegibilidade e a avaliação dos projetos. Neste sentido, foram apurados os respetivos *accruals*, através dos quatro modelos já identificados, bem como o cálculo de diversas variáveis independentes consideradas relevantes à modelação do problema.

3.3.1 Variáveis dependentes adotadas – Elegibilidade; Avaliação do projeto

As variáveis dependentes adotadas nesta investigação são a elegibilidade da candidatura e a avaliação do projeto que visam cumprir os objetivos específicos 1.2 e 2.1. As duas variáveis caracterizam-se por serem nominais face à intenção de distinguir o *status* das candidaturas. As Tabelas 14 e 15 refletem a distribuição das variáveis dependentes durante o período de 2012 a 2019.

Para a elegibilidade considera-se não elegível uma candidatura, quando esta não cumpre os pré-requisitos e o seu mérito é baixo ($n = 11.1136$). Neste sentido, cerca de metade das candidaturas

(49,67%) não são aprovadas, e apenas um terço (33,94%) são efetivamente apoiadas, uma vez que as candidaturas mesmo elegidas podem não vir a ser não apoiadas por falta de orçamento disponível ($n = 3.675$).

Tabela 14: Distribuição das candidaturas por elegibilidade entre os anos 2012 e 2019

	2012 ⁽¹⁾		2013		2015		2016		2017		2018		2019		Total	
	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Elegível	2.244	67,04	1.408	40,98	1.228	18,83	1.729	49,87	425	16,38	448	27,69	129	8,97	7.611	33,94
Não Apoiado	121	3,62	618	17,99	1.590	24,38	219	6,32	466	17,96	271	16,75	390	27,12	3.675	13,39
Não Elegível	982	29,34	1.410	41,03	3.704	56,79	1.519	43,81	1.703	65,66	899	55,56	919	63,91	11.136	49,67
Total	3.347	100,00	3.436	100,00	6.522	100,00	3.467	100,00	2.594	100,00	1.618	100,00	1.438	100,00	22.422	100,00

Notas: ⁽¹⁾ Não foi disponibilizada informação relativa a candidaturas do ano 2014.

Quanto à avaliação do projeto, considera-se que este foi bem-sucedido desde que sejam cumpridos os objetivos contratualizados em pelo menos 90%, ou seja, quando o mérito do projeto não é inferior em mais do que 10% ao mérito da candidatura contratualizada. No período entre 2012 e 2019, das 7.611 candidaturas apoiadas, apenas 40,56% tiveram sucesso, valor que poderá aumentar uma vez que 2.342 candidaturas aguardam a avaliação do mérito do projeto.

Tabela 15: Distribuição das candidaturas por avaliação do projeto entre os anos 2012 e 2019

	2012		2013		2015		2016		2017		2018		2019		Total	
	Nº	%	Nº	%												
Sucesso⁽¹⁾	324	14,44	201	14,28	661	53,83	1109	64,14	332	78,11	347	77,45	113	87,60	3087	40,56
Insucesso	401	17,87	396	28,13	560	45,60	618	35,74	92	21,65	99	22,10	16	12,40	2182	28,67
Por apurar⁽²⁾	1519	67,69	811	57,59	7	0,57	2	0,12	1	0,24	2	0,45	0	0	2342	30,77
Total	2244	100,00	1408	100,00	1228	100,00	1729	100,00	425	100,00	448	100,00	129	100,00	7.611	100,00

Notas: ⁽¹⁾ Quanto ao sucesso considerou-se uma *proxy* com valores de 2018 das vendas, e com o número dos empregados, considerando 10% de margem, ou seja, a partir de 90% foi considerado sucesso. ⁽²⁾ Aguardam a avaliação do mérito do projeto.

3.3.2 Variáveis independentes adotadas

No que concerne às variáveis independentes utilizadas neste estudo, estas dividem-se em cinco dimensões: demográfica, económico-financeira, recursos humanos, manipulação de resultados e projeto.

3.3.2.1 Manipulação de resultados

Para medir a MR, segundo Kasznick (1999), Kothari *et al.* (2005), e McNichols (2002), partiu-se das equações (8), (10) e (11), estimando-se os modelos de regressão linear através do método de

estimação OLS para cada ano, entre 2011 e 2018, e setores de atividade, considerados os cinco setores apresentados na Tabela 12.

É de realçar ainda, que para garantir a heterocedasticidade da amostra, foi utilizado como denominador comum o ativo total do ano anterior nos modelos de Kasznick e Kothari. Assim, a MR é obtida pelas equações apresentadas na Tabela 16, assumindo-se que a componente não discricionária dos *accruals* corresponde à parcela dos *accruals* totais (TA) não explicada por cada um desses modelos, ou seja, ao resíduo.

Tabela 16: Síntese dos Modelos de *Accruals* Não Discricionários – Regressão OLS

Modelo	Equação
Kasznick (1999)	$KAZ_{i,t} = \left[TA_{i,t} - \left(\hat{\alpha}_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \hat{\alpha}_2 \frac{(\Delta REV_{i,t} - \Delta AR_{i,t})}{A_{i,t-1}} + \hat{\alpha}_3 \frac{PPE_{i,t}}{A_{i,t-1}} + \hat{\alpha}_4 \frac{\Delta CFO_{i,t}}{A_{i,t-1}} \right) \right] \quad (14)$
Kothari <i>et al.</i> (2005)	$KOT_{i,t} = \left[TA_{i,t} - \left(\hat{\alpha}_0 + \alpha_1 \frac{1}{A_{i,t-1}} + \hat{\alpha}_2 \frac{(\Delta REV_{i,t} - \Delta AR_{i,t})}{A_{i,t-1}} + \hat{\alpha}_3 \frac{PPE_{i,t}}{A_{i,t-1}} + \hat{\alpha}_4 ROA_{i,t} \right) \right] \quad (15)$
McNichols (2002)	$MCN_{it} = [\Delta WC_{it} - (\hat{\alpha}_0 + \alpha_1 CFO_{it-1} + \hat{\alpha}_2 CFO_{it} + \hat{\alpha}_3 CFO_{it+1} + \hat{\alpha}_4 \Delta REV_{it+} + \hat{\alpha}_5 PPE_{it})] \quad (16)$

Já para medir a MR através do modelo desenvolvido por Beneish (1997), foi necessário fazer alterações às variáveis anteriormente apresentadas na Tabela 6, uma vez que não se dispunha de toda a informação necessária para o seu cálculo. Neste sentido foram consideradas *proxies* para as variáveis TATA, COGS, e SGA, sintetizadas na Tabela 17.

Tabela 17: Alterações às variáveis do modelo de Beneish

Variável	Fórmula
TATA	$\frac{(\Delta Working\ capital - \Delta cash - depreciation_t)}{Total\ Assets_t}$
COGS	<i>Material costs</i>
SGA	$Rev - (Material\ costs + Employes\ costs)$

Nota: *Material cost*: custos de material; *Employes costs*: custos dos empregados.

Para calcular a MR através dos quatro modelos, foi necessário calcular previamente as diferentes variáveis constantes das diferentes equações, a partir dos dados extraídos da base de dados Amadeus.

3.3.2.2 Outras dimensões de preditores da elegibilidade e da avaliação do projeto

A Tabela 18 apresenta, para além das variáveis independentes da MR, outras variáveis potenciais preditoras da elegibilidade e da avaliação do projeto, selecionadas de acordo com o especialista do IAPMEI e limitadas pelos dados disponibilizados por este organismo. Destaca-se a inclusão de variáveis

relativas aos gestores e *advisers*, pois estes têm todo o interesse em apresentar resultados aos sócios/acionistas, tal como preconiza a teoria da agência (Jense & Mackling, 1976).

Tabela 18: Descrição das variáveis independentes

Variável	Descrição	Tipo ⁽²⁾	Unidade de medida	Objetivo	Descritivas ⁽³⁾
Características Demográficas					
ANTIGUIDADE	Idade da empresa à data da candidatura	D	Meses	1.3/2.2	NR: 22370; M:138,0; DP:196,8
NUTSII_SEDE	Região	N	n.a.	1.3/2.2	NR: 22434; Mo: Norte (49,6%)
NUTSIII_SEDE	Concelho	N	n.a.	1.3/2.2	NR: 22422; Mo: Grande Porto (17,0%)
DISTRITO_SEDE	Distrito	N	n.a.	1.3/2.2	NR: 22422; Mo: Porto (23,6%)
DIMENSÃO	Dimensão da empresa	N	n.a.	1.3/2.2	NR: 22422; Mo: Micro (42,3%)
FORMA_LEGAL	Forma jurídica da empresa	N	n.a.	1.3/2.2	NR: 22422; Mo: LDA (60,4%)
SETOR_II	Setor da empresa	N	n.a.	1.3/2.2	NR: 22422; Mo: C (44,5%)
SETOR_III	Setor da empresa	N	n.a.	1.3/2.2	NR: 22422; Mo: 3 (50,8%)
Características Económico-financeiras					
LEVERAGE ⁽¹⁾	Alavancagem financeira	C	proporção	1.3/2.2	NR: 69947; M: 4,51; DP: 7,62
ROA ⁽¹⁾	Retorno do ativo	C	proporção	1.3/2.2	NR:71575; M:0,03; DP: 0,14
VAR.REV ⁽¹⁾	Variação anual das vendas	C	%	1.3/2.2	NR: 66623; M: 0,33; DP: 4,11
AT ⁽¹⁾	Ativo total	C	x1000 Euros	1.3/2.2	NR: 90796; M: 5941,7; DP: 16107,71
EXP_N1	Peso das vendas internacionais no ano N1	C	%	1.3/2.2	NR: 7382; M: 32,9; DP: 33,6
Características de RH					
NUM_EMP ⁽¹⁾	Nº de empregados por ano	D	pessoas	1.3/2.2	NR: 71334; M: 34,83; DP: 83,99
NUM_PREV_DM	Nº prévio de gestores	D	pessoas	1.3/2.2	NR: 13102; M: 3,0; DP: 3,92
NUM_CURR_DM	Nº atual de gestores	D	pessoas	2.2	NR: 21316; M: 3,0; DP: 3,31
NUM_PREV_AD	Nº prévio de consultores	D	pessoas	1.3/2.2	NR: 5911; M: 2,0; DP: 1,70
NUM_CURR_AD	Nº atual de consultores	D	pessoas	2.2	NR: 12579; M: 3,0; DP: 1,62
Manipulação de resultados					
KAZ ⁽¹⁾	MR: Modelo Kasznick	C	proporção	1.3/2.2	NR:66388; M: -0,03; DP: 0,80
ABS_KAZ ⁽¹⁾	Módulo do KAZ	C	proporção	1.3/2.2	NR: 66388; M:0,05; DP:0,66
SINAL_KAZ ⁽¹⁾	Sinal do KAZ	B	n.a.	1.3/2.2	NR: 66388; Mo:1 (55,6%)
KOT ⁽¹⁾	MR: Modelo Kothari	C	proporção	1.3/2.2	NR: 66889; M: -0,03; DP: 0,80
ABS_KOT ⁽¹⁾	Módulo do KOT	C	proporção	1.3/2.2	NR: 66889; M: 0,05; DP: 0,66
SINAL_KOT ⁽¹⁾	Sinal do KOT	B	n.a.	1.3/2.2	NR: 66889; Mo:1 (57,1%)
MCN ⁽¹⁾	MR: Modelo McNichols	C	proporção	1.3/2.2	NR:50911; M: -7,95; DP: 643,45
ABS_MCN ⁽¹⁾	Módulo do MCN	C	proporção	1.3/2.2	NR: 50911; M: 258,14; DP: 590,91
SINAL_MCN ⁽¹⁾	Sinal do MCN	B	n.a.	1.3/2.2	NR: 50911; Mo: 0 (51,2%)
MSCORE ⁽¹⁾	MR: Modelo Beneish	C	proporção	1.3/2.2	NR: 71632; M: -1,78; DP: 2,41
MANI_MSCORE ⁽¹⁾	Indicador de MR: sim/não	B	n.a.	1.3/2.2	NR: 71632; Mo: 1 (61,1%)

Variável	Descrição	Tipo ⁽²⁾	Unidade de medida	Objetivo	Descritivas ⁽³⁾
Características do projeto					
PROPÓSITO	Objetivo da candidatura	N	n.a.	1.3/2.2	NR: 22422; Mo: Qualificação (54,2%)
NUTSII_INV	Região	N	n.a.	1.3/2.2	NR: 22422; Mo: Norte (49,6%)
NUTSIII_INV	Concelho	N	n.a.	1.3/2.2	NR: 22422; Mo: Grande Porto (16,8%)
DISTRITO_INV	Distrito	N	n.a.	1.3/2.2	NR: 22422; Mo: Porto (23,4%)
DURAÇÃO_PROJETO	Duração do projeto candidato	D	meses	1.3/2.2	NR: 15538; M:23,0; DP: 6,2
NUM_EMP_PREPOS	Nº de postos de trabalho pré projeto	D	peessoas	1.3/2.2	NR: 13800; M: 13,0; DP:65,6
NUM_EMP_PREPOS	Nº de postos de trabalho pós projeto	D	peessoas	1.3/2.2	NR: 8319; M: 24,0; DP: 83,3
MERITO_CANDIDATURA	Score do mérito da candidatura	C	pontos	2.2	NR: 11286; M: 3,6; DP: 16,6
EXP_PP	Peso das vendas internacionais no ano pós projeto	C	%	2.2	NR: 10595; M: 50,0; DP: 28,2
INVESTIMENTO	Investimento total	C	X1000 euros	1.3/2.2	NR: 20932; M:122,6; DP:1511,2
INCENTIVO	Peso do incentivo na despesa elegível	C	%	2.2	NR: 10008; M: 70,0; DP:13,4
DESPESA_ELEGIVEL	Peso da despesa elegível no investimento	C	%	2.2	NR:14870; M:100; DP: 15,612

Notas: (1) Variáveis calculadas para os anos N-7 (identificada como N7) a N (identificada como N), sendo N o ano da candidatura. (2) B: (qualitativa) binária; N: qualitativa nominal; D: quantitativa discreta; C: quantitativa continua. (3) NR: número de respostas válidas; M: média; DP: desvio-padrão; Mo: moda; n.a.: não se aplica.

3.4 Técnicas de análise de dados

A qualidade da informação, cada vez mais ganha destaque, nomeadamente a informação contida numa base de dados de um estudo. Neste seguimento, constatou-se que a informação exportada da base de dados *Bureau van Dijk Amadeus* continha não respostas e *outliers*. Assim, as não respostas foram excluídas, o que não afetou a representatividade amostra uma vez que o número de candidaturas é elevado, e foram corrigidos os *outliers* tendo em conta a média e o desvio padrão de cada variável, considerando-se *outliers* sempre que o valor observado afasta-se em mais do que três desvios padrão da média. Nestes casos, substituiu-se os valores *outliers* pela média mais ou menos três vezes o desvio padrão, respetivamente, quando o valor se afasta para cima ou para baixo do padrão. Assim, diminui-se o impacto que os *outliers* possam ter na aplicação dos modelos preditivos (Jacinto, 2019).

Dada a informação disponibilizada, decidiu-se eleger o ano da data da candidatura como o ano de referência, ou seja, uma candidatura realizada no ano de 2017, terá como referência o ano de 2017, isto é, $N = 2017$, $N1 = 2016$, $N2 = 2015$ e assim sucessivamente. Consequentemente, os anos de 2020 e 2010 não foram totalmente analisados, uma vez que serviram como suporte de cálculo de certas variáveis, daí a sua incorporação na análise.

Quanto aos dados estes foram analisados recorrendo a diferentes técnicas de análise de dados. Inicialmente, para além dos modelos de regressão linear para estimar a MR, recorreu-se a técnicas de estatística descritiva, uni e bivariada, e a técnicas de inferência estatística, nomeadamente, aos testes de hipóteses, após verificados os seus pressupostos, para avaliar a significância das relações entre variáveis (Laureano, 2020). Depois, e atendendo à complexidade do problema, recorreu-se a técnicas de estatística multivariada, designadamente, a árvores de decisão com os algoritmos CART (*Classification and Regression Trees*) (Breiman *et al.*, 1984) e CHAID (*Chi-squared Automatic Interaction Detection*) (Kass, 1980). A Tabela 19 resume as principais técnicas, associando-as aos objetivos da investigação.

Tabela 19: Principais técnicas de análise de dados utilizadas

Técnicas	Objetivos
Técnicas de estatística descritiva univariada	Preparação de dados / 1.1
Análise e tratamento de <i>outliers</i>	Preparação de dados
Modelos de regressão linear múltipla	Preparação de dados
Técnicas de estatística descritiva bivariada	1.1 / 1.2 / 2.1 / 1.3 / 2.2
Testes de hipóteses – Pearson; teste t; Mann-Whitney	1.2 / 2.1
Árvores de decisão para classificação – CART e CHAID	1.3 / 2.2

Por fim, importa referir que o nível de significância (α), utilizado nos testes de hipóteses, é o de 0,05, ou seja, rejeita-se a hipótese nula (existindo, pois, diferenças ou relações significativas) quando a probabilidade de significância associada aos testes (*p-value*) for igual ou inferior ao α . No caso dos modelos preditivos criados, consideram-se bons modelos de classificação quando a percentagem de casos corretamente classificados, a sensibilidade e a especificidade forem, no mínimo, de 65%. Ainda para validar os modelos e avaliar a sua capacidade de generalização, utilizou-se uma amostra de teste (30% das candidaturas) para aplicar aos modelos criados com uma amostra de treino (70% das candidaturas) e determinou-se as métricas de avaliação dos modelos referidas anteriormente.

3.5 Modelo com Árvores de Decisão

As árvores de decisão podem ser classificadas em árvores de regressão, quando a variável de resposta é quantitativa, ou em árvores de classificação, quando a resposta é qualitativa (Pestana & Gageiro, 2014). Estas representam um conjunto de decisões, que se consubstanciam na definição de um conjunto de regras que se sucedem numa hierarquia simples e condicionada (Pereira *et al.*, 2010). Não obstante, o modelo com árvores de decisão caracteriza-se por representar regras facilmente interpretáveis, por possuir uma boa capacidade preditiva, e pela clareza dos seus resultados, tornando-se numa metodologia adequada para aplicação de análises preditivas (Rokach & Maimon, 2008). Adicionalmente, são técnicas que lidam bem com qualquer tipo e número de variáveis, não

respostas e *outliers*, não requerem a verificação de pressupostos e contemplam as possíveis interações entre as variáveis independentes (Constantino, 2020).

O processo de construção do modelo de árvores de decisão é composto por duas etapas: o treino e o teste, onde a amostra é dividida em dois grupos aleatórios, mutuamente exclusivos. No treino são estimados os parâmetros do modelo no sentido de encontrar o seu melhor resultado, que posteriormente será aplicado à amostra teste de forma a avaliar a sua precisão. Importa referir que na etapa treino são selecionados algoritmos de árvores de decisão, sendo o CART e o CHAID, os mais difundidos na literatura (Pestana & Gageiro, 2014).

Neste estudo, apenas foram abordadas árvores de classificação, uma vez que as variáveis elegibilidade e sucesso são qualitativas, construídas com base nos algoritmos CART e CHAID. O CART trata-se de um algoritmo não paramétrico, que cria árvores de decisões binárias, sendo que possui duas fases: a de crescimento e a de poda. Por sua vez o CHAID é um método exploratório que visa estudar as relações entre uma variável de resposta e um conjunto de variáveis explicativas que possam interagir entre si (Michalewicz *et al.*, 2006). Para cada um destes algoritmos foram utilizadas várias parametrizações apresentadas na Tabela 20, para o modelo preditivo da elegibilidade, e na Tabela 21, para o modelo preditivo da avaliação do projeto.

Os modelos obtidos através de algoritmos podem ser contruídos com recurso a uma árvore única (*single tree*), com recurso à técnica *bagging*, que visa melhorar a capacidade de generalização dos mesmos aumentando a estabilidade do modelo, ou através da técnica *boosting* que pretende melhorar a precisão do modelo na fase de treino. Após a definição do algoritmo, foi necessário determinar o número máximo de níveis da árvore, ou seja, a profundidade máxima da mesma, estipulado por defeito como cinco, uma árvore menos complexa e mais facilmente generalizável. Para além disso, procurou-se fixar o número mínimo de casos por Nó Pai (delimitado entre 2 e 8) e o número mínimo de casos por Nó Filho (delimitado por 1 e 2). Estas limitações ao crescimento das árvores visam evitar o sobreajustamento dos modelos aos dados de treino.

Tabela 20: Parametrização dos modelos preditivos selecionados para a elegibilidade

	Modelos									
	A	B	C	D	E	F	G	H	I	J
Algoritmo	CART	CART	CART	CART	CHAID	CHAID	CHAID	CHAID	CHAID	CHAID
Ensembles	-	<i>Boosting</i>	<i>Bagging</i>	-	-	<i>Boosting</i>	<i>Bagging</i>	<i>Bagging</i>	<i>Bagging</i>	-
Profundidade máxima	7	8	8	7	5	5	5	5	6	5
Número de casos Nó Pai	2	5	5	8	5	5	5	5	5	5
Número de casos Nó Filho	1	2	2	2	2	2	2	2	2	2
Custos de má classificação	Não	Não	Não	Não	Não	Não	Não	Não	Não	Não
Anos em análise	N1 a N7	N1 a N7	N1 a N7	N1 a N7	N1 a N7	N1 a N7	N1 a N7	N1 a N7	N1	N1

Em alguns dos modelos, nomeadamente nos modelos relacionados com o sucesso, foi necessário utilizar custos de má classificação de forma a equilibrar a qualidade entre o teste, com 70% dos casos, e o treino, com 30% dos casos, uma vez que os modelos gerados apresentavam uma maior qualidade no treino do que no teste. A utilização dos custos de má classificação permite, igualmente, pseudo equilibrar a amostra, aumentando a importância da categoria de elegibilidade menos representada e, assim, permitir que o modelo tenha boa capacidade preditiva, quer para as elegidas (sensibilidade), quer para as não elegidas (especificidade). Importa referir que inicialmente para a elegibilidade, o período definido para a construção dos modelos foi entre N1 e N7, ou seja, os anos anteriores à candidatura, e para o sucesso entre N a N7, o que posteriormente foi reduzido a apenas N1, e N e N1, respetivamente.

Tabela 21: Parâmetros dos modelos selecionados para a avaliação do projeto

	Modelos									
	A	B	C	D	E	F	G	H	I	J
Algoritmo	CART	CART	CART	CART	CART	CHAID	CHAID	CHAID	CHAID	CHAID
Ensembles	-	-	-	<i>Boosting</i>	<i>Bagging</i>	-	-	-	<i>Boosting</i>	<i>Bagging</i>
Profundidade máxima	7	7	9	9	9	5	7	9	9	9
Número de casos Nó Pai	2	2	5	5	5	5	5	5	5	5
Número de casos Nó Filho	1	1	2	2	2	2	2	2	2	2
Custos de má classificação	Não	Não	Não	Sim	Sim	Não	Sim	Sim	Sim	Sim
Anos em análise	N a N7	N e N1	N e N1	N e N1	N e N1	N e N1	N e N1	N e N1	N e N1	N e N1

A avaliação da qualidade dos modelos teve em consideração diversas métricas, sendo que algumas são baseadas na matriz de classificação (Delen *et al.* 2013), apresentada na Tabela 22.

Tabela 22: Matriz de classificação

Classe Observada	Classe Prevista (pelo modelo)	
	Não elegida/Insucesso	Elegida/Sucesso
Não elegida/Insucesso	VN (verdadeiro negativo)	FP (Falso positivo)
Elegida/Sucesso	FN (Falso negativo)	VP (Verdadeiro Positivo)

Em particular recorreu-se à percentagem de casos corretamente classificados (PCCC), à especificidade, que traduz a percentagem de casos corretamente classificados de não elegibilidade e de insucesso, e à sensibilidade, que representa a percentagem das candidaturas elegidas e candidaturas com sucesso que o modelo acertou. A Tabela 23 resume as fórmulas utilizadas para o cálculo destas métricas.

Tabela 23: Métricas de avaliação de qualidade dos modelos de classificação

Métrica	Fórmula
PCCC	$PCCC = \frac{VP + VN}{VP + VN + FP + FN}$
Especificidade	$E = \frac{VN}{VN + FP}$
Sensibilidade	$S = \frac{VP}{VP + FN}$

Ainda foi utilizada uma outra métrica, a AUC (*Area Under the ROC Curve*), que reflete a precisão global do modelo, considerando quer os valores de especificidade, quer os valores de sensibilidade. Assim, quanto maior o valor da AUC, maior será o poder do modelo, sendo que o seu valor pode variar entre 0, que corresponde a um modelo com previsões 100% erradas, e 1 que corresponde a um modelo com previsões 100% corretas (Mingers, 1989).

4. Resultados e Discussão

Este capítulo visa dar sequência à metodologia abordada no capítulo anterior, evidenciando os resultados das análises efetuadas e sua interpretação, e discussão para cada um dos objetivos definidos.

4.1 Avaliação da manipulação de resultados nas candidaturas

A análise da MR por quatro modelos diferentes (Tabela 24) permite concluir que as empresas tendem a manipular os seus resultados. Mais concretamente, e em relação aos modelos McNichols (MCN) e Beneish (MSCORE), estes avaliaram a MR no sentido da qualidade dos *accruals*, isto é, quanto maior for o seu afastamento em relação a zero, menor será a qualidade dos *accruals*. Os resultados sugerem, ainda, que a MR ocorre em ambos os sentidos, ou seja, tanto no sentido positivo, quando se oculta da declaração financeira uma situação financeira mais degradada, como no sentido negativo, quando a MR ocorre de forma a minimizar o imposto sobre o rendimento a pagar (Carmo *et al.*, 2010; Lisboa, 2016). Assim, a Tabela 24 apresenta as medidas descritivas para os quatro modelos estudados, quer para cada ano, quer para um conjunto de anos.

Relativamente aos modelos Kasznick (KAZ) e Kothari (KOT), que visam quantificar a MR, constata-se que em termos médios existe uma tendência para manipulação nula, como esperado. No entanto, existem empresas que manipulam intensamente no sentido de diminuir os resultados, mínimos entre -5,11 e -13,82, para o KAZ, e entre -5,31 e -13,66 para o KOT, e empresas que manipulam de forma a aumentar os resultados, máximos entre 3,71 e 8,99, e entre 2,72 e 8,83, respetivamente. Ressalva-se que em termos médios e medianos, para estes dois modelos, a tendência é para não existir MR nas demonstrações financeiras.

No que respeita ao MCN, verifica-se que em todos os anos, a MR em termos médios é negativa, e ronda os 8 pontos, e em termos medianos, é relativamente próximo de zero, com exceção no ano N1 e N7 (3,29 e -6,90, respetivamente). Estes indicadores variam entre -8.914,21 para o ano N4 e 9.553,05 para o ano N1. Já o modelo MSCORE apresenta, em média, valores negativos e que rondam 1,80 pontos, revelando uma tendência clara que existe MR, uma vez que a média é superior ao valor de referência, -2,22 pontos, a partir do qual as empresas são consideradas manipuladoras. Ressalva-se ainda que o mesmo se verifica em termos medianos. No entanto, este indicador varia entre -9,99, referindo-se a uma empresa que claramente não manipula, uma vez que é bastante inferior ao valor de referência do modelo, e um máximo de 14,98 pontos.

Neste contexto, cumpre realçar que os valores obtidos encontram-se em consonância com estudos anteriores, destacando-se Im e Nam (2019), Ittonen *et al.* (2013), Gill-de-Albornoz e Illuena (2005) e Talab *et al.* (2017) para os modelos KAZ, KOT, MCN e MSCORE, respetivamente.

Tabela 24: Medidas descritivas das variáveis de manipulação de resultados

Modelos	Nº	Média	Desvio padrão	Mínimo	Mediana	Máximo
KAZ_N	17.255	-0,04	0,79	-12,19	0,04	7,71
KAZ_N1	13.879	-0,03	0,85	-13,82	0,07	6,80
KAZ_N2	11.269	-0,03	0,80	-11,74	0,06	8,99
KAZ_N3	10.771	-0,02	0,78	-11,00	0,07	6,80
KAZ_N4	6.297	-0,02	0,74	-8,35	0,05	7,42
KAZ_N5	3.853	-0,01	0,81	-12,89	0,07	3,71
KAZ_N6	2.067	-0,05	0,79	-7,87	0,03	4,20
KAZ_N7	997	-0,05	0,69	-5,11	0,01	3,98
KAZ_TOTAL	-	-0,03	0,80	-13,91	0,05	8,98
KOT_N	17.382	-0,03	0,80	-11,92	0,06	7,57
KOT_N1	13.971	-0,03	0,85	-13,66	0,08	6,61
KOT_N2	11.368	-0,02	0,81	-11,25	0,08	8,83
KOT_N3	10.847	-0,02	0,79	-11,09	0,08	6,61
KOT_N4	6.348	-0,01	0,75	-8,65	0,06	7,27
KOT_N5	3.884	-0,01	0,76	-8,98	0,08	3,58
KOT_N6	2.089	-0,04	0,79	-7,88	0,04	3,43
KOT_N7	1.000	-0,05	0,68	-5,31	0,03	2,72
KOT_TOTAL	-	-0,03	0,80	-13,66	0,07	8,83
MCN_N	12.733	-8,32	644,13	-8477,84	-0,50	8771,84
MCN_N1	11.341	-8,28	597,44	-7914,21	3,29	9553,05
MCN_N2	10.231	-7,13	632,81	-8819,76	-0,27	8891,11
MCN_N3	7.654	-7,49	684,79	-8680,28	1,77	8842,09
MCN_N4	4.588	-8,31	620,54	-8914,21	-1,09	8856,60
MCN_N5	2.676	-8,17	708,28	-8129,33	-0,41	8792,67
MCN_N6	1.341	-8,85	710,12	-8659,21	-0,31	8000,49
MCN_N7	347	-7,77	898,19	-7612,25	-6,90	8187,68
MCN_TOTAL	-	-7,94	643,45	-8914,21	0,90	9553,05
MSCORE_N	18.784	-1,77	2,50	-9,99	-1,81	14,94
MSCORE_N1	14.969	-1,72	2,46	-9,90	-1,79	14,96
MSCORE_N2	12.168	-1,81	2,47	-9,92	-1,82	14,98
MSCORE_N3	11.524	-1,78	2,32	-9,93	-1,79	14,98
MSCORE_N4	6.782	-1,80	2,33	-9,96	-1,79	14,94
MSCORE_N5	4.148	-1,82	2,27	-9,95	-1,81	14,26
MSCORE_N6	2.208	-1,83	2,10	-9,60	-1,78	14,10
MSCORE_N7	1.049	-1,84	1,88	-8,95	-1,84	13,39
MSCORE_TOTAL	-	-1,78	2,41	-9,98	-1,80	14,97

Relativamente às correlações entre os diferentes modelos, a Tabela 25 evidencia que dentro de cada modelo, para KAZ e para KOT, não existem relações relevantes entre este indicador nos diferentes anos ($R < 0,1$), o que contrasta com a situação apresentada para os modelos MCN e MSCORE, em que as correlações já são fracas a moderadas. Assim, é relevante considerarem-se estes indicadores em diferentes anos.

Tabela 25: Matriz de correlações de *Pearson* na manipulação de resultados

	KAZ_N	KAZ_N1	KAZ_N4	KAZ_N7	KOT_N	KOT_N1	KOT_N4	KOT_N7	MCN_N	MCN_N1	MCN_N4	MCN_N7	MSCORE_N	MSCORE_N1	MSCORE_N4	MSCORE_N7
KAZ_N	1	-0,058	-0,020	-0,013	0,994	-0,057	-0,023	-0,007	0,108	-0,010	-0,010	0,018	-0,054	0,019	0,015	-0,002
KAZ_N1		1	-0,015	-0,090	-0,059	0,995	-0,012	-0,093	0,017	0,098	-0,013	0,137	-0,024	-0,082	0,040	-0,010
KAZ_N4			1	0,033	-0,019	-0,008	0,996	0,049	-0,016	0,008	0,116	-0,063	0,001	-0,015	-0,074	0,087
KAZ_N7				1	-0,020	-0,083	0,053	0,996	0,024	-0,012	0,044	0,067	-0,001	-0,055	-0,036	-0,128
KOT_N					1	-0,057	-0,022	-0,016	0,103	-0,011	-0,012	-0,008	-0,052	0,024	0,022	0,007
KOT_N1						1	-0,007	-0,092	0,020	0,096	-0,014	0,171	-0,014	-0,080	0,042	-0,035
KOT_N4							1	0,067	-0,017	0,012	0,117	-0,057	0,006	-0,025	-0,070	0,089
KOT_N7								1	0,023	-0,009	0,054	0,066	-0,011	-0,068	-0,023	-0,122
MCN_N									1	-0,043	-0,074	0,268	-0,081	0,015	0,004	0,067
MCN_N1										1	0,071	-0,273	-0,025	-0,095	0,006	-0,069
MCN_N4											1	-0,530	-0,018	-0,030	-0,083	-0,025
MCN_N7												1	0,001	0,041	-0,031	-0,031
MSCORE_N													1	0,343	0,214	0,231
MSCORE_N1														1	0,231	0,306
MSCORE_N4															1	0,225
MSCORE_N7																1

Notas: O menor número de observações 317; Negrito: correlação significativa para um nível de significância de 0,05.

Para o modelo MCN, e mais concretamente no ano N7, existe uma correlação fraca entre este e os anos N (R=0,268) e N1 (R=-0,273), e moderada com o ano N4 (R=-0,530). No entanto, estas ocorrem em sentidos opostos, uma vez que a relação com N é positivamente correlacionada, e com os restantes esta correlação tem sinal negativo. Referente ao MSCORE, constata-se que as relações entre os diferentes anos são significativas, sendo estas de intensidade fraca e positivas (coeficientes entre 0,214 e 0,343). No entanto, apesar de existir alguma correlação, sendo esta relativamente fraca, optou-se por considerar nas análises variáveis relativas a todos os anos.

Numa outra perspetiva, e analisando a relação que existe entre os diferentes modelos e os diferentes anos, os modelos KAZ e KOT são praticamente equivalentes (correlação positiva quase perfeita). De facto, os modelos são muito semelhantes entre si, nomeadamente na sua fórmula de cálculo, como explicitado na Tabela 16, uma vez que entre estes dois modelos apenas é alterada um indicador. Ainda assim, e de acordo com estudos anteriores (Fisher *et al.*, 2019), decidiu-se manter os dois modelos com o objetivo de esclarecer se estes modelos de MR são relevantes para a explicação da elegibilidade e do sucesso nas candidaturas. Enfatiza-se, ainda, que esta opção não inviabiliza a aplicação da metodologia utilizada, já que os modelos preditivos baseados em árvores de decisão não colocam em causa o princípio da multicolinearidade, ou seja, as relações entre as variáveis independentes não constituem uma limitação *à priori* para a aplicação destas técnicas.

4.2 Relação entre a manipulação de resultados e a elegibilidade das candidaturas

Tendo em vista o segundo objetivo, que visa caracterizar a relação entre a MR e a elegibilidade, procedeu-se à análise da relação entre estas duas variáveis. As Tabelas de 26 a Tabela 29 apresentam a comparação das medidas descritivas e os resultados dos testes t e *Mann-Whitney*.

A análise destes resultados permite evidenciar, para o modelo KAZ, que nos anos N1 e N3, existem diferenças significativas, quer em termos médios, quer em termos de distribuição, entre as candidaturas elegidas e não elegidas. De facto, nestes anos as candidaturas elegidas apresentam, em média, uma manipulação em sentido de redução dos resultados significativamente maior do que nas não elegidas. Nos restantes anos, constata-se que as diferenças não são significativas. Quanto ao modelo de Kothari as conclusões revelam resultados semelhantes. De facto, o seu comportamento semelhante, quanto à elegibilidade, corrobora a correlação quase perfeita que existe entre os dois modelos, como referido anteriormente.

Tabela 26: Distribuição do Kasznick por elegibilidade da candidatura

MR	Elegibilidade	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste T; Mann-Whitney
KAZ_N1	Não Elegível	7199	-0,02	0,87	-13,82	0,07	6,80	$t(13877) = 2,219; p = 0,027$ $MW_Z = -2,169; p = 0,030$
	Elegível	6680	-0,05	0,83	-11,44	0,06	6,58	
KAZ_N2	Não Elegível	6065	-0,01	0,83	-11,74	0,06	8,99	$t(11208) = 1,813; p = 0,070$ $MW_Z = -1,215; p = 0,225$
	Elegível	5204	-0,04	0,77	-9,43	0,06	4,70	
KAZ_N3	Não Elegível	5756	0,00	0,81	-11,00	0,08	6,80	$t(10743) = 2,678; p = 0,007$ $MW_Z = -4,631; p < 0,001$
	Elegível	5015	-0,04	0,74	-11,00	0,05	4,37	
KAZ_N4	Não Elegível	3343	-0,02	0,78	-8,35	0,05	6,24	$t(6295) = -0,565; p = 0,572$ $MW_Z = -0,454; p = 0,650$
	Elegível	2954	-0,01	0,68	-6,24	0,05	7,42	
KAZ_N5	Não Elegível	2331	-0,01	0,86	-12,89	0,07	3,37	$t(3851) = -0,164; p = 0,869$ $MW_Z = -1,052; p = 0,293$
	Elegível	1522	-0,01	0,72	-7,06	0,06	3,71	
KAZ_N6	Não Elegível	1171	-0,05	0,86	-7,13	0,04	4,20	$t(2064) = -0,259; p = 0,796$ $MW_Z = -0,914; p = 0,361$
	Elegível	896	-0,04	0,68	-7,87	0,02	2,49	
KAZ_N7	Não Elegível	593	-0,03	0,73	-5,11	0,01	3,98	$t(995) = 0,957; p = 0,339$ $MW_Z = -1,132; p = 0,258$
	Elegível	404	-0,07	0,62	-3,73	0,02	2,19	

Notas: Nº: número de candidatura D.P.: desvio padrão.

Tabela 27: Distribuição do Kothari por elegibilidade da candidatura

MR	Elegibilidade	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste T; Mann-Whitney
KOT_N1	Não Elegível	7263	-0,01	0,87	-13,66	0,09	6,61	$t(13969) = 1,993; p = 0,046$ $MW_Z = -2,147; p = 0,032$
	Elegível	6708	-0,04	0,83	-11,44	0,07	6,45	
KOT_N2	Não Elegível	6128	-0,01	0,83	-11,25	0,08	7,57	$t(11260) = 1,746; p = 0,081$ $MW_Z = -1,131; p = 0,258$
	Elegível	5240	-0,04	0,78	-8,93	0,07	8,83	
KOT_N3	Não Elegível	5801	0,00	0,82	-11,08	0,09	6,61	$t(10815) = 2,323; p = 0,020$ $MW_Z = -4,555; p < 0,001$
	Elegível	5046	-0,04	0,75	-11,09	0,06	4,90	
KOT_N4	Não Elegível	3385	-0,02	0,81	-8,65	0,07	5,90	$t(6335) = -0,344; p = 0,731$ $MW_Z = -0,794; p = 0,427$
	Elegível	2963	-0,01	0,68	-6,60	0,06	7,27	
KOT_N5	Não Elegível	2354	-0,01	0,77	-8,98	0,08	3,27	$t(3882) = 0,723; p = 0,470$ $MW_Z = -1,430; p = 0,153$
	Elegível	1530	-0,02	0,73	-7,14	0,06	3,58	
KOT_N6	Não Elegível	1186	-0,04	0,87	-7,17	0,05	3,41	$t(2084) = -0,070; p = 0,945$ $MW_Z = -1,343; p = 0,179$
	Elegível	903	-0,04	0,68	-7,88	0,04	3,43	
KOT_N7	Não Elegível	599	-0,02	0,73	-5,31	0,04	2,72	$t(998) = 1,576; p = 0,115$ $MW_Z = -1,871; p = 0,061$
	Elegível	401	-0,09	0,60	-3,76	0,02	2,20	

Notas: Nº: número de candidatura D.P.: desvio padrão.

Em relação ao modelo de McNichols, os anos N1, N2, N3 e N7 apresentam diferenças significativas entre as candidaturas eleitas e não eleitas. No entanto, para os anos N1 e N3, estas só são diferentes significativamente em termos de distribuição, o que contraria os anos N2 e N7, que apresentam diferenças significativas em termos médios da MR. Os restantes anos não apresentam diferenças significativas, nem em relação à média, nem em relação à distribuição.

Tabela 28: Distribuição do McNichols por elegibilidade da candidatura

MR	Elegibilidade	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste T; Mann-Whitney
MCN_N1	Não Elegível	5929	-2,11	512,79	-7070,48	4,29	7612,76	$t(10038) = 1,136; p = 0,256$ $MW_Z = -2,281; p = 0,023$
	Elegível	5412	-15,03	678,13	-7914,21	1,36	9553,05	
MCN_N2	Não Elegível	5367	5,99	546,41	-7882,37	-0,08	8386,05	$t(9063) = 2,175; p = 0,030$ $MW_Z = -1,243; p = 0,014$
	Elegível	4864	-21,61	715,94	-8819,76	-0,50	8891,11	
MCN_N3	Não Elegível	4021	0,00	584,72	-6850,21	4,24	8842,09	$t(6693) = 0,993; p = 0,321$ $MW_Z = -2,475; p = 0,013$
	Elegível	3633	-15,79	780,74	-8680,28	-2,53	8411,21	
MCN_N4	Não Elegível	2568	-0,34	578,11	-8914,21	-0,63	8856,60	$t(3995) = 0,964; p = 0,335$ $MW_Z = -1,279; p = 0,201$
	Elegível	2020	-18,45	670,63	-7271,83	-2,88	8367,61	
MCN_N5	Não Elegível	1564	10,59	681,62	-8129,33	1,45	6708,61	$t(2261) = 1,602; p = 0,109$ $MW_Z = -1,120; p = 0,263$
	Elegível	1112	-34,56	743,67	-8041,64	-3,12	8792,67	
MCN_N6	Não Elegível	773	-7,69	792,17	-8659,21	-1,33	8000,49	$t(1339) = 0,070; p = 0,944$ $MW_Z = -0,388; p = 0,698$
	Elegível	568	-10,44	580,80	-6436,41	1,99	3110,68	
MCN_N7	Não Elegível	210	-31,79	1042,85	-7612,25	-14,60	8187,68	$t(345) = -0,616; p = 0,538$ $MW_Z = -2,496; p = 0,013$
	Elegível	137	29,04	615,55	-3586,23	18,07	3024,22	

Notas: Nº: número de candidatura D.P.: desvio padrão.

Para o MSCORE, todos os anos, com a exceção de N6, constata-se diferenças significativas, quer em termos médios, quer em termos de distribuição, entre as candidaturas elegidas e não elegidas. Mais concretamente, nestes anos as candidaturas não elegidas apresentam um *score* significativamente mais negativo do que nas elegidas, contudo, em termos médios, ambos os grupos amostrais, apresentam valores superiores a -2,22 pontos, o que significa que, em média, estas empresas consideram-se manipuladoras.

Tabela 29: Distribuição do M-Score por elegibilidade da candidatura

MR	Elegibilidade	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste t; Mann-Whitney
MSCORE_N1	Não Elegível	7887	-1,80	2,59	-9,90	-1,87	14,96	$t(14965) = -4,264; p < 0,001$ $MW_Z = -5,073; p < 0,001$
	Elegível	7082	-1,63	2,30	-9,57	-1,72	14,55	
MSCORE_N2	Não Elegível	6643	-1,92	2,58	-9,92	-1,92	14,98	$t(12089) = -5,428; p < 0,001$ $MW_Z = -6,205; p < 0,001$
	Elegível	5525	-1,68	2,32	-9,88	-1,74	14,96	
MSCORE_N3	Não Elegível	6219	-1,89	2,42	-9,90	-1,88	13,60	$t(11475) = -5,613; p < 0,001$ $MW_Z = -5,684; p < 0,001$
	Elegível	5305	-1,65	2,20	-9,93	-1,71	14,98	
MSCORE_N4	Não Elegível	3639	-1,89	2,44	-9,96	-1,86	14,66	$t(6769) = -3,197; p = 0,001$ $MW_Z = -2,993; p = 0,003$
	Elegível	3143	-1,71	2,19	-9,60	-1,74	14,94	
MSCORE_N5	Não Elegível	2523	-1,90	2,40	-9,95	-1,88	14,06	$t(3856) = -3,080; p = 0,002$ $MW_Z = -3,711; p < 0,001$
	Elegível	1625	-1,69	2,03	-9,70	-1,70	14,26	
MSCORE_N6	Não Elegível	1261	-1,87	2,20	-9,60	-1,81	14,10	$t(2142) = -1,119; p = 0,263$ $MW_Z = -1,486; p = 0,137$
	Elegível	947	-1,77	1,96	-9,12	-1,75	12,16	
MSCORE_N7	Não Elegível	639	-1,96	1,79	-7,45	-1,92	10,34	$t(1047) = -2,588; p = 0,010$ $MW_Z = -2,719; p = 0,007$
	Elegível	410	-1,66	2,00	-8,95	-1,75	13,39	

Notas: Nº: número de candidatura D.P.: desvio padrão.

Com estes resultados, pode considerar-se que a hipótese I de investigação encontra-se verificada pelo modelo do Beneish (MSCORE), onde se constata que, com exceção do ano N6, em todos os anos existem diferenças significativas, quer em termos médios, quer em termos de distribuição da MR, entre as candidaturas elegidas e não elegidas. Pelos restantes modelos, apenas em alguns anos se identificam diferenças entre os dois tipos de candidaturas, pelo que se considera que os resultados são inconclusivos. Assim, considera-se a hipótese I verificada essencialmente devido ao cálculo da MR na perspetiva de Beneish, isto é, atendo à qualidade dos *accruals*.

4.3 Modelo preditivo para elegibilidade das candidaturas

Atendendo ao terceiro objetivo específico estimaram-se modelos preditivos da elegibilidade das candidaturas tendo por base dois algoritmos de árvores de decisão para classificação. Os resultados dos diferentes modelos estimados, de acordo com o ponto 3.5 da metodologia, são apresentados na Tabela 30, quer para a amostra treino (70%), quer para a amostra teste (30%). Os resultados evidenciam que os modelos baseados no algoritmo CART apresentam pior qualidade do que os modelos obtidos com o algoritmo CHAID. De facto o melhor modelo obtido com CART apresenta uma PCCC no treino de apenas 73,09%, enquanto que no CHAID, atingiu-se um PCCC no treino de 94,65%, embora este último resultado possa sugerir algum sobreajustamento no modelo. No entanto, no teste, os melhores resultados foram obtidos no modelo G com um PCCC de 78,73%, e é neste valor que se deve focar a escolha do modelo, uma vez que o teste, consiste na aplicação do modelo criado às novas candidaturas.

Assim, e perante os seis modelos criados com o algoritmo CHAID, o melhor é o G, que requer acesso a informação dos sete anos anteriores à candidatura, quando existem empresas que não têm atividade há sete anos. Assim, optou-se por escolher o modelo H. Este apenas tem em consideração o ano anterior à candidatura, o ano N1, e cuja qualidade no teste é muito ligeiramente inferior ao modelo G (PCCC: 78,01%; E: 80,23%; S: 75,90%; AUC: 0,869).

Tabela 30: Resultados dos modelos preditivos para a elegibilidade

		Modelo									
		A	B	C	D	E	F	G	H	I	J
Treino (70%)	PCCC	72,49%	75,90%	75,57%	73,09%	81,81%	94,65%	86,16%	85,15%	88,40%	80,70%
	E	72,44%	76,90%	90,67%	70,13%	82,79%	94,95%	87,80%	87,07%	90,57%	78,29%
	S	72,54%	74,89%	60,42%	76,07%	80,82%	94,36%	84,51%	83,23%	86,22%	83,12%
	AUC	0,772	0,845	0,846	0,815	0,912	0,988	0,943	0,935	0,954	0,903
		CART				CHAID					
Teste (30%)	PCCC	72,14%	72,59%	73,23%	71,18%	77,37%	75,94%	78,73%	78,01%	77,96%	77,07%
	E	72,12%	73,52%	88,62%	67,08%	77,76%	76,08%	80,81%	80,23%	80,08%	74,41%
	S	72,16%	71,70%	58,62%	75,06%	77,00%	75,81%	76,77%	75,90%	75,95%	79,60%
	AUC	0,763	0,799	0,823	0,798	0,869	0,847	0,873	0,869	0,866	0,868
		CART				CHAID					

Notas: E: especificidade; S: sensibilidade.

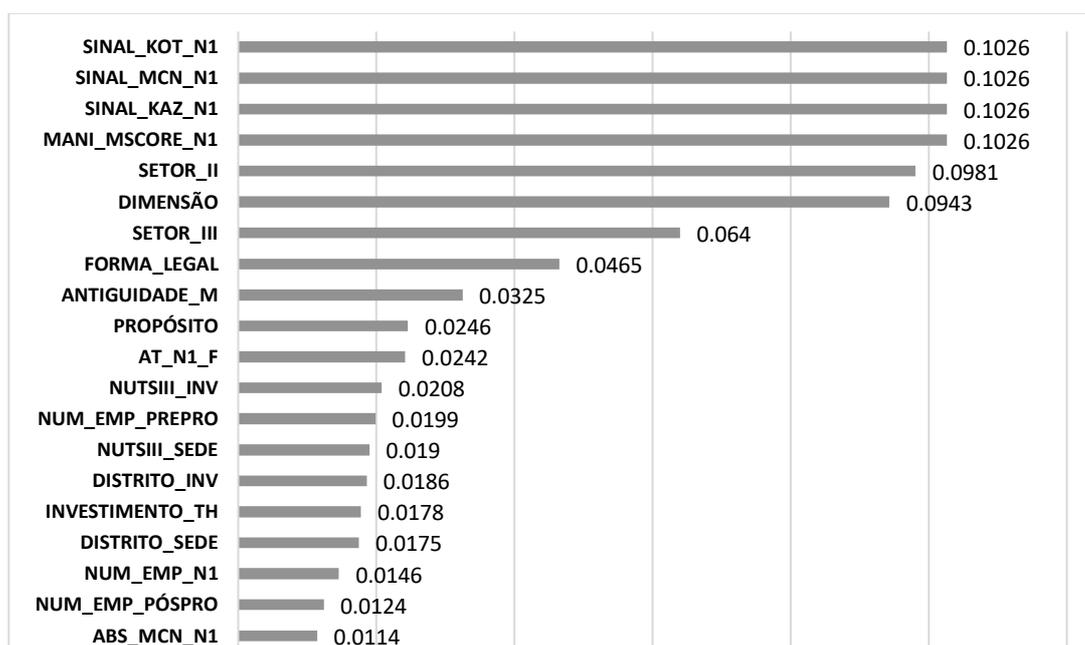
Adicionalmente, verifica-se no modelo G, que tem em consideração os últimos sete anos, que entre as quatro variáveis mais importantes, três referem-se a anos muito anteriores à candidatura (N7 e N6), como se verifica na Tabela 31. Destaca-se, ainda e numa perspetiva geral, que entre as variáveis do *top 4* surgem com frequência as que refletem o sinal da manipulação, ou seja, o sentido em que esta ocorre, aumentar ou diminuir os resultados (e.g. SINAL_KOT_N1), e não aquelas que procuram quantificar a intensidade da manipulação (MSCORE_N6).

Tabela 31: Importância das variáveis dos modelos preditivos para a elegibilidade

Modelos	Importância das variáveis – TOP 4			
A	MSCORE_N6	DURAÇÃO_PROJETO	KOT_N7	LEVERAGE_N3
B	SINAL_KAZ_N3	SINAL_KOT_N5	SINAL_KOT_N4	MANI_MSCORE_N3
C	SINAL_KOT_N5	MANI_MSCORE_N6	SINAL_KAZ_N5	SINAL_KOT_N3
D	SINAL_MCN_N2	SINAL_KOT_N1	SINAL_KAZ_N1	SINAL_KAZ_N6
E	SINAL_MCN_N4	SETOR_III	PROPÓSITO	AT_N1
F	MANI_MSCORE_N6	MANI_MSCORE_N7	SINAL_KOT_N4	SINAL_KOT_N2
G	SINAL_MCN_N7	SINAL_KOT_N7	SINAL_KOT_N1	SINAL_KOT_N6
H	SINAL_KOT_N1	SINAL_MCN_N1	SINAL_KAZ_N1	MANI_SCORE_N1
I	SINAL_KOT_N1	SINAL_KAZ_N1	MANI_MSCORE_N1	SINAL_MCN_N1
J	DURAÇÃO_PROJETO	EXP_N1	SETOR_III	DIMENSÃO

Mais concretamente, e através da análise da Figura 4, verifica-se que as variáveis que apresentam maior importância preditiva para o modelo H pertencem à dimensão da MR, e procuram esclarecer quanto ao sentido da manipulação. Esta dimensão, apenas volta a surgir, na 20ª posição, com o valor absoluto do modelo do McNichols para o ano N1 (ABS_MCN_N1).

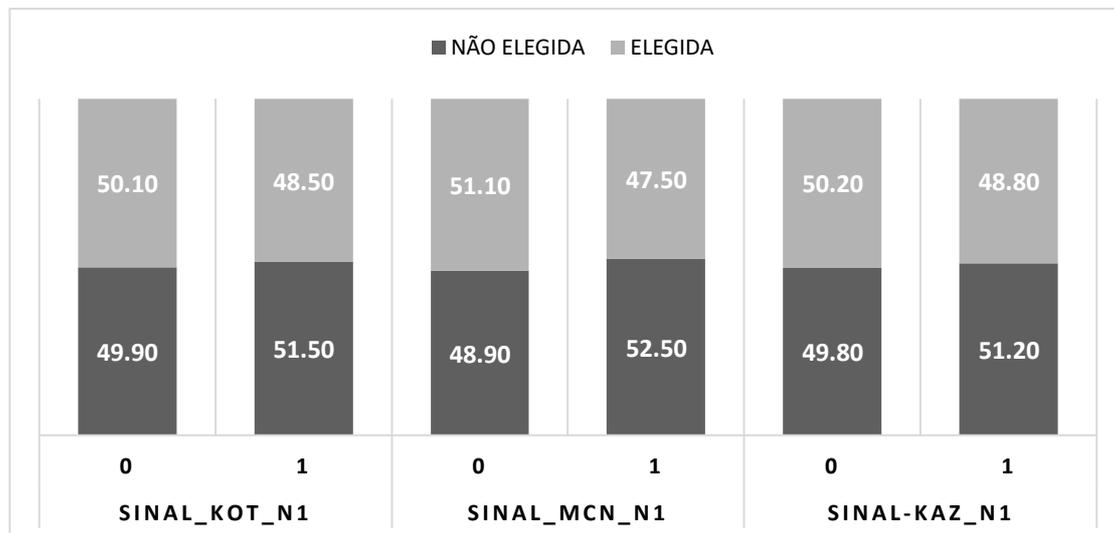
Figura 4: Importância das variáveis predictoras da elegibilidade, relativas ao modelo H



As características demográficas ocupam o *top 5*, através do SETOR_II, que apresenta um nível de importância de 0,098, enquanto que o PROPÓSITO, variável referente às características do projeto ocupa um lugar mais baixo (10º) com apenas 0,024 de importância. Os dados de avaliação financeira da empresa, apenas surgem na 11ª posição (AT_N1), sendo a última a dimensão a surgir, são os recursos humanos, apenas no 18º lugar (NUM_EMP_N1).

A elegibilidade prevista de uma candidatura, quando relacionada com as três variáveis com maior importância preditiva para o modelo H, apresenta distribuição semelhante nos três casos. De facto, os três indicadores, apresentados na Figura 5, refletem o sentido da manipulação, ou seja, tomam valor 0, se esta ocorre no sentido de diminuição dos resultados (sinal negativo), e 1, se esta ocorre no sentido de aumento dos resultados (sinal positivo). Nos três casos, verifica-se que existe maior probabilidade, ainda que ligeira, de uma empresa, quando o sentido da manipulação é negativo, não ser elegida.

Figura 5: Distribuição da elegibilidade prevista para as três variáveis mais importantes



Deste modo, verifica-se, que uma empresa que se candidate a fundos europeus, terá maior probabilidade de ser financiada se, de forma a cumprir os pré-requisitos ou aumentar o mérito da candidatura, opte por manipular os resultados no sentido negativo no ano N1, sendo esta probabilidade ligeiramente maior para o modelo MCN (51,10%). De forma oposta, se a candidatura refletir resultados superiores à realidade para o ano N1, esta terá maior probabilidade de não ser elegida (51,50% para o KOT, 52,50% para o KAZ e 51,20% para o MCN), acabando por não beneficiar dos apoios.

De forma a identificar perfis de candidaturas elegidas, não elegidas ou de difícil classificação, uma vez que o melhor modelo (H) resulta de uma combinação de modelos (*bagging*), recorre-se, a título de exemplo, a nós terminais com elevada confiança (maior 80%) e algum suporte (N >20) do modelo J.

Perfis de candidaturas com elevada propensão a ser não elegidas traduzidos em regras proposicionais:

- se (DURAÇÃO_PROJETO >18 e DURAÇÃO_PROJETO <=22), e EXP_N1 <= 6,005, e (SETOR_II = 3 ou SETOR_II = 4) e, MANI_MSCORE_N1 = 0, então candidatura = não elegida (Nó 144: suporte = 22 candidaturas; confiança = 1,0);
- se (DURAÇÃO_PROJETO >14 e DURAÇÃO_PROJETO <=18), NUM_EMP_POSPRO > 9 e, (PROPÓSITO = IDT ou PROPÓSITO = QUALIDADE), então candidatura = não elegida (Nó 62: suporte = 22 candidaturas; confiança = 1,0).

Perfis de candidaturas com elevada propensão a ser elegidas:

- se (DURAÇÃO_PROJETO = 10 ou DURAÇÃO_PROJETO = 11), INVESTIMENTO <= 20 e, NUM_EMP_N1 > 82, então candidatura = elegida (Nó 46: suporte = 62 candidaturas; confiança = 0,984);
- se DURAÇÃO_PROJETO <= 10 e, INVESTIMENTO < 450,798 e, AT_N1 > 1208,626 e, (DIMENSÃO = MICRO ou DIMENSÃO = MÉDIA) e, (MANI_MSCORE_N1 = 1 ou MANI_MSCORE_N1 = 0) e, então candidatura = elegida (Nó 207: suporte = 41 candidaturas; confiança = 0,829).

Perfil de candidaturas de difícil classificação, correspondente ao nó folha 240 da árvore, se (DURAÇÃO_PROJETO >=22 ou DURAÇÃO_PROJETO <=23) e, EXP_N1 >= 2,867 e EXP_N1 <= 12,100 e, SETOR_III <=4 e, DIMENSÃO = MICRO e, NUM_EMP_N1 <= 13 e, então candidatura = não elegida (suporte = 77 candidaturas; confiança = 0,507).

Adicionalmente, importa referir que a análise das diferentes regras, que resultam dos nós folha da árvore de decisão, vêm confirmar uma das vantagens da utilização das árvores de decisão, que é o efeito da interação das variáveis preditoras na decisão de classificar uma candidatura em elegível ou não elegível. Por exemplo, verifica-se que a não MR no ano N1 (MANI_MSCORE_N1 = 0) tanto está associado à classificação de candidaturas como não elegíveis (nó 144) como elegíveis (nó 207), fazendo-se a diferença pelo facto desta variável interagir com outras (conjunção de condições). Muitas interações entre muitas diferentes variáveis, levando a classificações diferentes, revela a complexidade do problema e a adequabilidade das árvores de decisão para modelar este problema (Delen *et al.*, 2013; West *et al.*, 2016).

Tendo como suporte os estudos identificados na literatura, pode-se afirmar que os resultados obtidos para o modelo preditivo da elegibilidade a fundos europeus encontram-se em consonância com os mesmos (Fernandes, 2011). Face ao exposto, o modelo defendido por Beneish revelou-se um método de cálculo de MR adequado à análise desta problemática (Fernandes & Laureano, 2019). Neste contexto, pode considerar-se que o objetivo de criação de um modelo preditivo para a elegibilidade,

tendo por base variáveis independentes relacionadas com os quatro modelos de MR considerados e de mais quatro dimensões, foi atingido.

4.4 Relação entre a manipulação de resultados e a avaliação do projeto

Tendo em vista o quarto objetivo, que procura caracterizar a relação entre a MR e a avaliação do projeto, procedeu-se à análise da relação entre estas duas variáveis. As Tabelas 32 a 35 apresentam a comparação das medidas descritivas e os resultados dos testes *t* e *Mann-Whitney*.

A análise destes resultados permite evidenciar, para o modelo KAZ, no ano N, N1 e N3, existem diferenças significativas entre as candidaturas com sucesso e insucesso. De facto, para os anos N e N3, estas apenas apresentam diferenças significativamente em termos de distribuição, sendo que o ano N1, reflete diferenças, e significativas, em termos médios da MR. Nos restantes anos, constata-se que as diferenças não são significativas. Quanto ao modelo de Kothari as conclusões apresentam-se idênticas.

Tabela 32: Distribuição do Kasznick por avaliação do projeto

MR	A.P.	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste T; Mann-Whitney
KAZ_N	Insucesso	1902	-0,09	0,72	-8,65	0,01	3,01	$t(4225) = -1,307; p = 0,191$ $MW_Z = -2,337; p = 0,019$
	Sucesso	2395	-0,06	0,80	-8,50	0,04	3,42	
KAZ_N1	Insucesso	1549	-0,05	0,75	-8,35	0,06	3,29	$t(3613) = 2,498; p = 0,013$ $MW_Z = -1,440; p = 0,150$
	Sucesso	2084	-0,12	0,95	-11,44	0,04	3,81	
KAZ_N2	Insucesso	1194	-0,07	0,80	-8,38	0,04	3,64	$t(3031) = -0,894; p = 0,372$ $MW_Z = -1,666; p = 0,096$
	Sucesso	1839	-0,05	0,80	-9,43	0,07	4,23	
KAZ_N3	Insucesso	1151	-0,08	0,60	-4,93	0,01	1,92	$t(2894) = -0,005; p = 0,996$ $MW_Z = -2,086; p = 0,037$
	Sucesso	1765	-0,08	0,85	-11,00	0,04	2,38	
KAZ_N4	Insucesso	664	-0,04	0,63	-6,24	0,02	1,97	$t(1564) = -0,472; p = 0,637$ $MW_Z = -0,607; p = 0,544$
	Sucesso	1270	-0,02	0,75	-5,21	0,05	7,42	
KAZ_N5	Insucesso	173	-0,03	0,67	-5,13	0,04	1,09	$t(698) = 0,044; p = 0,965$ $MW_Z = -0,522; p = 0,602$
	Sucesso	527	-0,04	0,77	-3,86	0,05	3,37	
KAZ_N6	Insucesso	94	-0,07	0,63	-3,33	0,00	1,44	$t(406) = 0,094; p = 0,925$ $MW_Z = -0,044; p = 0,965$
	Sucesso	314	-0,07	0,80	-7,87	0,01	1,72	
KAZ_N7	Insucesso	11	0,16	0,41	-0,38	0,10	1,20	$t(100) = 1,453; p = 0,149$ $MW_Z = -1,634; p = 0,102$
	Sucesso	91	-0,15	0,68	-3,73	-0,02	1,41	

Notas: A.P: Avaliação do projeto; Nº: número de candidatura D.P.: desvio padrão.

Tabela 33: Distribuição do *Kothari* por avaliação do projeto

MR	A.P.	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste T; Mann-Whitney
KOT_N	Insucesso	1896	-0,07	0,72	-8,77	0,02	2,97	t(4259) = -1,396; p = 0,163 MW_Z = -3,061; p = 0,002
	Sucesso	2421	-0,04	0,82	-8,41	0,06	3,36	
KOT_N1	Insucesso	1544	-0,04	0,75	-8,41	0,07	3,05	t(3620) = 2,300; p = 0,022 MW_Z = -1,089; p = 0,276
	Sucesso	2098	-0,11	0,95	-11,44	0,05	3,58	
KOT_N2	Insucesso	1199	-0,06	0,80	-8,41	0,05	3,67	t(3061) = -0,450; p = 0,653 MW_Z = -1,474; p = 0,141
	Sucesso	1864	-0,05	0,84	-8,93	0,08	8,83	
KOT_N3	Insucesso	1153	-0,07	0,62	-4,93	0,03	4,90	t(2902) = 0,289; p = 0,773 MW_Z = -2,006; p = 0,045
	Sucesso	1784	-0,07	0,86	-11,09	0,05	2,38	
KOT_N4	Insucesso	665	-0,03	0,63	-6,60	0,02	1,75	t(1933) = -0,091; p = 0,927 MW_Z = -0,445; p = 0,656
	Sucesso	1270	-0,03	0,73	-5,29	0,06	7,27	
KOT_N5	Insucesso	174	-0,02	0,77	-7,14	0,07	1,34	t(706) = 0,644; p = 0,520 MW_Z = -1,108; p = 0,268
	Sucesso	534	-0,07	0,80	-4,14	0,04	2,94	
KOT_N6	Insucesso	94	-0,06	0,64	-3,32	0,01	1,46	t(409) = -0,053; p = 0,958 MW_Z = -0,276; p = 0,783
	Sucesso	317	-0,06	0,83	-7,88	0,04	3,43	
KOT_N7	Insucesso	11	0,18	0,43	-0,36	0,13	1,24	t(98) = 1,756; p = 0,082 MW_Z = -1,867; p = 0,062
	Sucesso	89	-0,19	0,66	-3,76	-0,06	0,85	

Notas: A.P.: Avaliação do projeto; Nº: número de candidatura D.P.: desvio padrão.

Para o modelo de McNichols, existem diferenças significativas em termos de médios e em termos de distribuição no ano N2, enquanto que para o ano N3 estas são apenas significativas em termos de distribuição. Os restantes anos as diferenças não se constituem significativas, nem em relação à média, nem em relação à distribuição do sucesso das candidaturas.

Tabela 34: Distribuição do McNichols por avaliação do projeto

MR	A.P.	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste t; Mann-Whitney
MCN_N	Insucesso	1631	-16,88	870,15	-7914,21	-12,04	8522,41	t(3606) = -0,405; p = 0,685 MW_Z = -1,467; p = 0,142
	Sucesso	1977	-4,81	907,23	-8477,84	-2,79	8187,68	
MCN_N1	Insucesso	1292	-37,10	875,75	-7914,21	-18,62	8682,19	t(3091) = -0,786; p = 0,432 MW_Z = -1,679; p = 0,093
	Sucesso	1801	-13,49	783,77	-6859,48	-2,56	9553,05	
MCN_N2	Insucesso	1168	-65,08	956,18	-8819,76	-19,70	8830,48	t(2042) = -1,975; p = 0,048 MW_Z = -2,887; p = 0,004
	Sucesso	1713	0,05	722,86	-5969,64	1,82	8891,11	
MCN_N3	Insucesso	860	20,74	1094,93	-7837,71	0,26	8411,21	t(1402) = 1,840; p = 0,066 MW_Z = -2,867; p = 0,004
	Sucesso	1361	-58,10	775,97	-8680,28	-10,56	6749,34	
MCN_N4	Insucesso	351	-49,21	1022,14	-7271,83	7,18	8367,61	t(486) = -0,212; p = 0,804 MW_Z = -1,178; p = 0,239
	Sucesso	784	-36,62	662,68	-5969,64	-13,10	5483,10	
MCN_N5	Insucesso	121	-112,07	1346,22	-8041,64	-12,80	6275,59	t(139) = -0,273; p = 0,785 MW_Z = -0,040; p = 0,968
	Sucesso	380	-77,35	664,65	-5047,36	-13,48	4308,20	
MCN_N6	Insucesso	51	-86,23	596,14	-2161,65	-61,84	1585,73	t(201) = -0,600; p = 0,549 MW_Z = -1,168; p = 0,243
	Sucesso	152	-16,76	751,03	-6436,41	-29,73	3110,68	
MCN_N7	Insucesso	5	-260,37	756,92	-1579,58	-96,61	252,85	t(32) = -0,869; p = 0,392 MW_Z = -0,754; p = 0,451
	Sucesso	29	97,09	862,32	-2797,04	39,51	3024,22	

Notas: A.P.: Avaliação do projeto; Nº: número de candidatura D.P.: desvio padrão.

Adicionalmente, e em relação ao MSCORE, a maioria dos anos apresenta diferenças significativas. Mais precisamente, para os anos N e N3, a distribuição da MR em relação ao sucesso das candidaturas difere entre os dois grupos populacionais. Acrescenta-se ainda, que para os anos N2, N5, e N6, verificam-se diferenças significativas quer ao nível do valor médio, quer ao nível da distribuição da manipulação. Para os restantes anos, quer a média, quer a distribuição da MR não são significativamente diferentes, ou seja, não existe evidência estatística de que a MR observada no grupo populacional das candidaturas com sucesso seja diferente, em termos médios, e em termos de distribuição, ao do grupo populacional das candidaturas com insucesso. Contudo, em termos médios, e para todos os anos, tantos os projetos com insucesso, como os projetos com sucesso, apresentam valores superiores a -2,22 pontos, o que significa que, em média, estas empresas consideram-se manipuladoras.

Tabela 35: Distribuição do M-Score por avaliação do projeto

MR	A.P.	Nº	Média	D.P.	Mínimo	Mediana	Máximo	Teste t; Mann-Whitney
MSCORE_N	Insucesso	1980	-1,67	2,03	-8,68	-1,81	14,46	$t(4543) = 1,357; p = 0,201$ $MW_Z = -2,888; p = 0,045$
	Sucesso	2620	-1,69	2,40	-9,99	-1,72	14,06	
MSCORE_N1	Insucesso	1598	-1,63	2,12	-8,82	-1,77	14,26	$t(3695) = -0,219; p = 0,826$ $MW_Z = -0,581; p = 0,561$
	Sucesso	2245	-1,61	2,45	-9,57	-1,72	14,06	
MSCORE_N2	Insucesso	1241	-1,63	2,24	-9,51	-1,81	14,96	$t(2761) = 3,062; p = 0,009$ $MW_Z = -3,040; p = 0,008$
	Sucesso	1977	-1,70	2,39	-9,84	-1,73	14,41	
MSCORE_N3	Insucesso	1204	-1,77	1,89	-8,66	-1,81	13,30	$t(2880) = -1,643; p = 0,100$ $MW_Z = -2,615; p = 0,009$
	Sucesso	1896	-1,65	2,27	-9,83	-1,64	14,35	
MSCORE_N4	Insucesso	696	-1,76	1,89	-9,46	-1,79	12,46	$t(1639) = -0,450; p = 0,653$ $MW_Z = -0,123; p = 0,902$
	Sucesso	1360	-1,72	2,27	-9,57	-1,72	14,94	
MSCORE_N5	Insucesso	180	-1,28	1,71	-4,94	-1,35	8,18	$t(366) = 3,057; p = 0,002$ $MW_Z = -3,254; p = 0,001$
	Sucesso	575	-1,76	2,12	-9,70	-1,80	10,96	
MSCORE_N6	Insucesso	97	-1,31	2,01	-6,81	-1,54	10,00	$t(425) = 2,783; p = 0,006$ $MW_Z = -2,022; p = 0,043$
	Sucesso	330	-1,88	1,71	-8,59	-1,79	6,17	
MSCORE_N7	Insucesso	12	-1,30	1,68	-5,28	-1,03	0,84	$t(101) = 1,340; p = 0,183$ $MW_Z = -1,552; p = 0,121$
	Sucesso	91	-1,86	1,32	-5,26	-1,85	1,92	

Notas: A.P.: Avaliação do projeto; Nº: número de candidatura D.P.: desvio padrão.

Nesta sequência, a hipótese II é considerada verificada pelo modelo do Beneish (MSCORE), onde na maioria dos anos, ou seja, com exceção do ano N1, N4, e N7, existem diferenças significativas, quer em termos médios, quer em termos de distribuição da MR entre as candidaturas com sucesso e insucesso. Pelos restantes modelos, apenas em alguns anos se identificam diferenças entre os dois tipos de candidaturas, pelo que se considera que os resultados são inconclusivos. A hipótese II é assim validada, essencialmente, devido ao cálculo da qualidade dos accruals, através do modelo defendido por Beneish.

4.5 Modelo preditivo para a avaliação do projeto

Os diferentes modelos estimados, de acordo com a secção 5 da metodologia, são apresentados na Tabela 36, quer para a amostra treino, quer para a amostra teste. Os resultados evidenciam, que os modelos CART apresentam ligeiramente pior qualidade do que os modelos com o algoritmo CHAID, uma vez que o melhor modelo obtido com CART apresenta um PCCC no treino de apenas 83,39%, enquanto que no CHAID, atingiu-se um PCCC no treino de 85,54%. Assim, e diante os cinco modelos criados com o algoritmo CHAID, o melhor é o modelo I (PCCC: 76,18%; E: 70,85%; 81,89%; AUC: 0,841).

Tabela 36: Resultados dos modelos preditivos para a avaliação do projeto

		Modelo									
		A	B	C	D	E	F	G	H	I	J
Treino (70%)		CART					CHAID				
	PCCC	66,05%	64,68%	67,56%	83,39%	78,85%	71,51%	73,62%	74,00%	84,67%	85,54%
	E	75,54%	65,69%	63,20%	86,01%	81,74%	78,70%	71,33%	71,39%	88,77%	89,90%
	S	56,37%	63,67%	71,90%	80,64%	75,90%	64,34%	75,90%	76,52%	80,67%	81,23%
	AUC	0,671	0,664	0,706	0,858	0,829	0,799	0,799	0,820	0,859	0,867
Teste (30%)	PCCC	66,57%	67,77%	68,74%	81,78%	80,73%	70,43%	73,53%	76,18%	84,93%	84,84%
	E	75,50%	67,11%	61,87%	83,68%	81,18%	77,31%	70,97%	70,85%	88,92%	85,31%
	S	58,00%	68,44%	75,76%	80,09%	80,31%	63,51%	76,24%	81,89%	80,71%	84,39%
	AUC	0,677	0,699	0,696	0,851	0,836	0,779	0,779	0,841	0,860	0,857

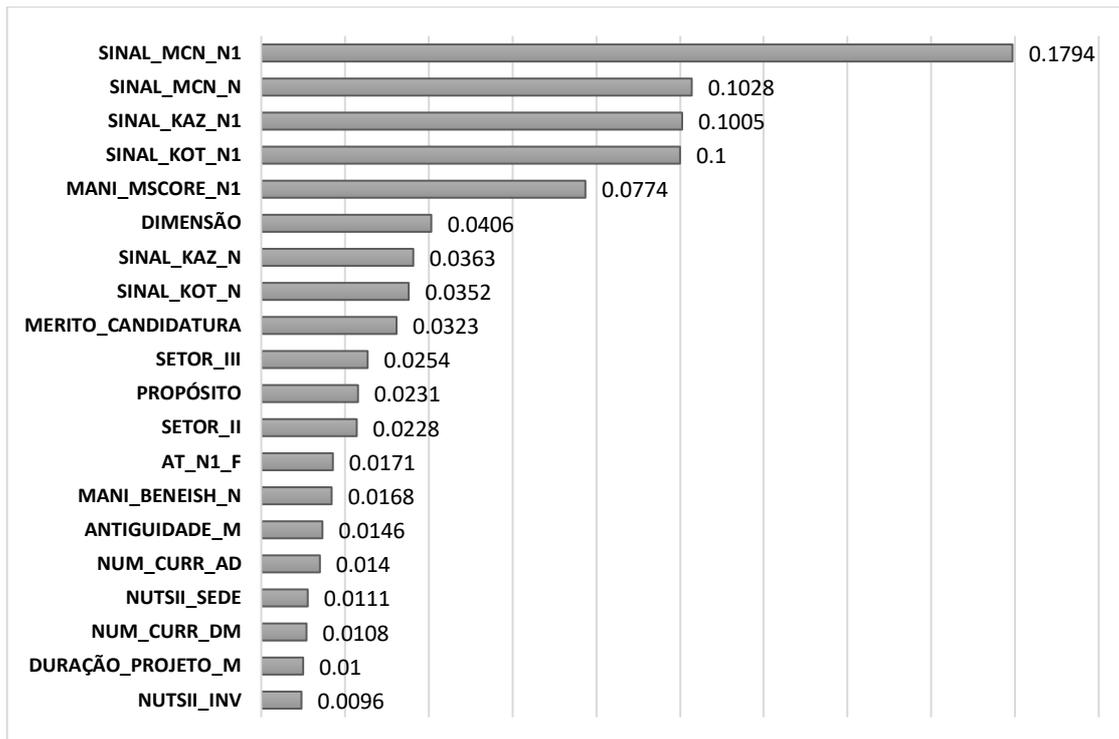
Adicionalmente, e numa perspetiva geral, verifica-se que entre as variáveis apresentam maior importância na predição dos modelos, as que surgem com maior frequência são o ativo total do ano N1 (AT_N1), o mérito da candidatura (MERITO_CANDIDATURA) e o sinal da MR nos diferentes modelos (e.g. SINAL_MCN_N1 e SINAL_KOT_N1)

Tabela 37: Importância das variáveis dos modelos preditivos para a avaliação do projeto

Modelos	Importância das variáveis – TOP 4			
A	MSCORE_N6	DURAÇÃO DO PROJETO	KOT_N7	LEVERAGE_N3
B	KOT_N2	KAZ_N2	VAR.REV_N2	LEVERAGE_N6
C	AT_N1	INCENTIVO	MERITO_CANDIDATURA	SETOR_II
D	MERITO_CANDIDATURA	AT_N1	INCENTIVO	ABS_MCN_N
E	SINAL_MCN_N1	SINAL_MCN_N	SINAL_KAZ_N1	SINAL_KOT_N1
F	AT_N1	FORMA_LEGAL	MERITO_CANDIDATURA	NUM_PREV_DM
G	AT_N1	PROPÓSITO	NUM_EMP_POSPRO	SINAL_MCN_N1
H	ROA_N	MERITO_CANDIDATURA	SINAL_KAZ_N	PROPOSITO
I	SINAL_MCN_N1	SINAL_MCN_N	SINAL_KAZ_N1	SINAL_KOT_N1
J	SINAL_MCN_N1	SINAL_KAZ_N1	SINAL_KOT_N1	SINAL_MCN_N

Através da análise da importância preditiva das variáveis para o modelo I, obteve-se a Figura 6, onde se encontram refletidos os primeiros 20 indicadores. Neste sentido, constata-se que as cinco variáveis que apresentam maior importância preditiva para o modelo I, pertencem à dimensão da MR, sendo que o SINAL_MCN_N1 é a que mais se destaca, ao contrário do que se verificou na elegibilidade em que os diferentes modelos apresentam a mesma importância no *top 4*.

Figura 6: Importância das variáveis predictoras da avaliação do projeto, relativas ao modelo I

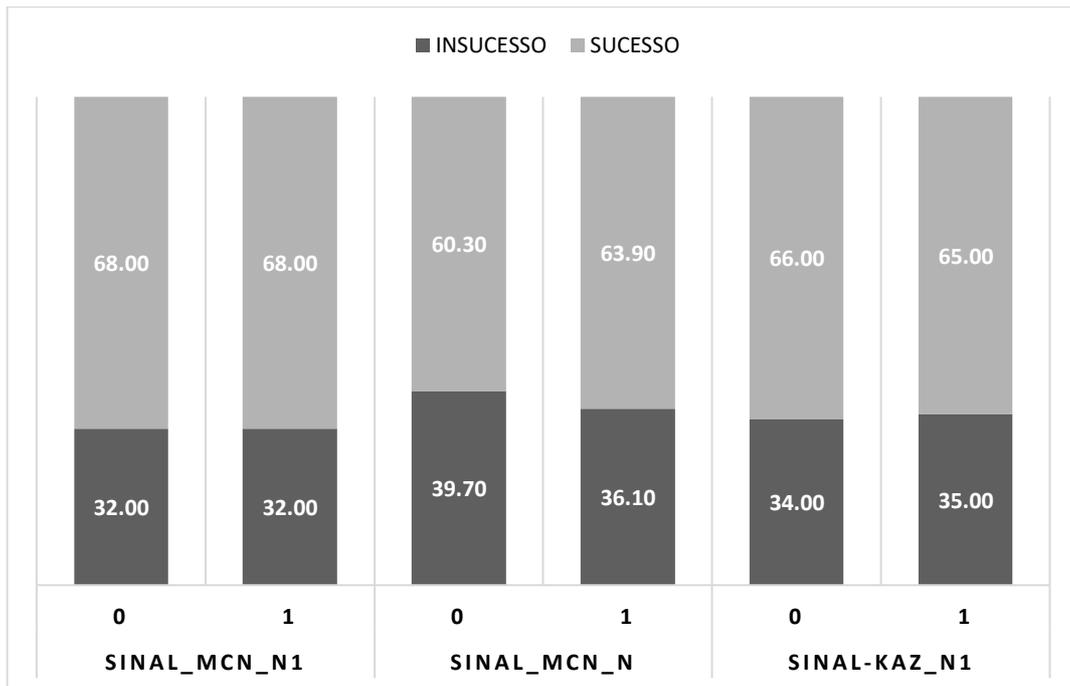


As restantes dimensões, todas elas se encontraram representadas neste *top 20*, sendo que a dimensão demográfica surge logo de seguida, representada pela variável DIMENSÃO (6ª posição) e as características do projeto exibem-se em 9ª lugar (MERITO_CANDIDATURA). Os dados de avaliação financeira da empresa, apenas surgem na 13ª posição (AT_N1), um pouco mais atrás quando comparado com o modelo preditivo eleito para a elegibilidade. Por último, surge a dimensão que procura caracterizar os recursos humanos da empresa, através da variável NUM_CURR_AD.

Os indicadores, apresentados na Figura 7, refletem o sentido em que ocorre a manipulação, ou seja, tomam valor 0 se esta ocorre no sentido de diminuir os resultados (sinal negativo), e 1 se esta ocorre no sentido de aumentar os resultados (sinal positivo). Ao analisar, o sucesso através dos três indicadores com maior importância preditiva para o modelo I, comprova-se que independente de qual seja a variável escolhida, e o sentido da MR, existe, sempre, maior probabilidade de a candidatura ser bem-sucedida. Por outro lado, e comparando apenas o modelo de McNichols, a probabilidade de ter

sucesso aumenta, se forem alvo de manipulação as demonstrações referentes ao ano anterior à candidatura (68% para MCN_N1). No entanto, como o efeito destas 3 variáveis de manipulação não é apenas considerada de forma isolada, mas também como efeito de interação com outras variáveis, leva a que a previsão da avaliação final do projeto não seja apenas sucesso, como a figura sugere, mas também insucesso.

Figura 7: Distribuição do sucesso prevista para as três variáveis mais importantes



Para a identificar perfis de candidaturas elegidas/não elegidas ou de difícil classificação, uma vez que o melhor modelo (I) resulta de uma combinação de modelos (*boosting*), recorre-se a título de exemplo, a nós terminais com elevada confiança (maior 80%) e algum suporte ($N > 20$) do modelo G.

Perfis de candidaturas com elevada propensão a ter insucesso traduzidos em regras proposicionais:

- se ($AF_N1 > 1313,036$ e $AF_N1 \leq 3631,818$) e, $SETOR_III \leq 3$, e ($NUM_EMP_POSPRO > 9$ e $NUM_EMP_POSPRO \leq 34$) e, $PROPOSITO = INOVAÇÃO$ e, $LEVERAGE_N1 > 1,781$, então candidatura = insucesso (nó 123: suporte = 38 candidaturas; confiança = 1,0);
- se ($AF_N1 > 3631,818$ e $AF_N1 \leq 9365,815$) e, $INCENTIVO > 65,025$ e, $SINAL_MCN_N1 = 1$ e, $ABS_MCN_N \leq 509,642$ e, ($SETOR_II = C$ ou $SETOR_II = M$ ou $SETOR_II = Q$ ou $SETOR_II = R$) e, ($DIMENSÃO = MICRO$ ou $DIMENSÃO = PEQUENA$) e, $NUM_EMP_N \leq 57$ e, $INVESTIMENTO \leq 2940$ então candidatura = insucesso (nó 215: suporte = 34 candidaturas; confiança = 1,0).

Perfis de candidaturas com elevada propensão a ter sucesso:

- se ($AF_N1 > 118$ e $AF_N1 \leq 492$) e, $NUM_EMP_PREPRO \leq 30$ e, $DURAÇÃO_PROJETO > 22$ e, ($SETOR_II = A$ ou $SETOR_II = B$ ou $SETOR_II = C$ ou $SETOR_II = E$ ou $SETOR_II = G$ ou $SETOR_II = J$ ou $SETOR_II = M$ ou $SETOR_II = N$ ou $SETOR_II = N$) e, $ABS_KAZ_N1 > 0,483$ e, $NUM_CURR_AD \leq 1$, então candidatura = sucesso (nó 162: suporte = 164 candidaturas; confiança = 0,945);
- se ($AF_N1 > 118,492$ e $AF_N1 \leq 616,446$) e $NUM_EMP_POSPRO \leq 139$ e, $SINAL_MCN_N = 1$ e, $NUM_PREV_DM \leq 2$ e, $VAR.REV_N1 \leq 0,147$ e, $EXP_N1 > 14,877$ e, então candidatura = sucesso (nó 164: suporte = 28 candidaturas; confiança = 0,1).

Perfis de candidaturas de difícil classificação: se ($AF_N1 > 1313,036$ e $AF_N1 \leq 3631,818$) e $SETOR_III \leq 3$ e, $NUM_EMP_POSPRO > 34$ e, ($ABS_MCN_N1 > 67,970$ e $ABS_MCN_N1 \leq 479,425$) e, $MSCORE_N1 \leq -2,933$ e, então candidatura = insucesso (nó 126: suporte = 31 candidaturas; confiança = 0,636).

Mais uma vez, é possível testemunhar o efeito da interação das variáveis preditoras na decisão de classificar um projeto em sucesso ou insucesso. Por exemplo, a MR no sentido positivo, ou seja, no aumento dos resultados ($SINAL_MCN_N = 1$) tanto está associado a projetos com insucesso (nó 215) como com sucesso (nó 164), sendo apenas possível esta distinção, uma vez que a variável não atua sozinha, mas sim em conjunto com as restantes (conjunto de condições). Nesta perspetiva, importa realçar a complexidade na previsão deste problema, e os bons resultados obtidos, o que não seria possível de alcançar com uma regressão normal.

Adicionalmente, é de se realçar que este modelo não teve apenas em consideração variáveis relacionadas com a MR, aliás foram incluídas variáveis de cinco dimensões diferentes. Isto significa, que a obtenção de um modelo promissor, que prevê corretamente cerca de 84% dos sucessos e fracassos, deve-se em grande parte à existência de inúmeras variáveis que representam as mais variadas características, ou seja, é indiciado que quanto maior for o número de diferentes variáveis maior serão as chances de que se venha a obter um bom modelo.

De facto, este estudo vai ao encontro com os resultados apresentados anteriormente, por Fernandes (2011), onde também foi alvo de estudo a ausência de eficácia na atribuição de fundos. O estudo anterior, modelou a avaliação do projeto através de regressão logística e de redes neuronais, utilizando apenas variáveis económico-financeiras e as características do projeto, tendo obtido um modelo que apenas prevê corretamente 78% dos casos. Neste contexto, considera-se que a presente dissertação, através da utilização de árvores de decisão e da inclusão de variáveis que caracterizam a MR, que foram de longe as mais importantes, foi possível melhorar a capacidade preditiva do modelo em aproximadamente 6%. Neste sentido, pode considerar-se que o objetivo de criação de um modelo preditivo para a avaliação do projeto, tendo por base variáveis independentes relacionadas com os quatro modelos de MR considerados e de mais quatro dimensões, foi atingido.

5. Conclusões

Em Portugal, o processo da atribuição de fundos europeus às empresas tem-se revelado pouco eficiente, uma vez que estudos realizados demonstram que estes têm sido deficientemente aproveitados. De forma a aumentar a eficácia e a eficiência na atribuição destes incentivos, é imperativo que estes sejam alocados a empresas que garantam, verdadeiramente, o cumprimento do seu contrato com o Estado Português. Nesta sequência, é necessário fornecer aos organismos intermédios ferramentas que permitam ajudar a avaliar as candidaturas, de forma a que os resultados contratualizados com as empresas, cujas candidaturas foram elegíveis sejam, efetivamente, concretizados.

Perante este problema, o presente estudo procura identificar uma relação entre a MR e o sucesso das candidaturas na atribuição de incentivos europeus, mais especificamente, através da avaliação do impacto da MR, quer no sucesso das candidaturas, quer no sucesso dos projetos financiados. Uma vez que a avaliação de candidaturas é um processo complexo, recorreu-se a técnicas ditas de *data mining* para a criação de modelos preditivos capazes de determinar perfis de candidaturas com sucesso, isto é, candidaturas que tenham elevada probabilidade em cumprir os objetivos contratualizados previamente.

Adicionalmente, esta investigação permite concluir que a MR pode ser avaliada quanto à sua intensidade e à sua qualidade através de diferentes modelos baseados em *accruals*. Na realidade, foi o modelo de Beneish, modelo que procura avaliar a qualidade dos *accruals*, que mais enaltece as diferenças de MR entre as candidaturas elegidas e não elegidas e, também, entre os projetos com sucesso e com insucesso. Por outro lado, comprovou-se que os modelos Kasznick e Kothari apresentam muitas semelhanças entre si. Na verdade, a sua correlação constatou-se quase perfeita, pelo que a presença de um poderá ser substituída por outro. Tal não se concretizou neste estudo, uma vez que o intuito era entender que variáveis mais contribuem para o sucesso dos projetos, o que para isso, e dada a complexidade do problema, é necessária a inclusão de diversas variáveis, de distintas dimensões do processo de candidatura, e das empresas candidatas aos incentivos financeiros. Assim, e através de modelos preditivos de classificação, concluiu-se que é a dimensão da MR que mais contribui para a previsão do sucesso dos projetos, mais concretamente, a variável que avalia o sentido da MR do modelo de McNichols.

Com base nos resultados obtidos, os objetivos foram concretizados e as duas hipóteses de investigação formuladas verificadas. De facto, os resultados comprovam em grande parte que a manipulação de resultados influencia o sucesso da candidatura a incentivos financeiros, tendo-se verificado diferentes níveis de manipulação entre candidaturas elegíveis e não elegíveis e, também, que a manipulação de resultados influencia a avaliação final dos projetos de investimento aprovados, tendo-se verificado diferentes níveis de manipulação consoante o projeto tenha sucesso ou insucesso.

Deste modo, o estudo permite responder à questão de investigação, uma vez que se constatou que os modelos baseados em *accruals* constituem um bom método para calcular os *accruals* discricionários, e indiretamente a manipulação de resultados. Mais concretamente, os modelos que avaliam a qualidade (modelo de McNichols e Beneish) apresentaram melhores resultados, quer na explicitação das diferenças entre os grupos populacionais estudados (não elegidas/elegidas e insucesso/sucesso), quer nos modelos preditivos, tanto no estudo da elegibilidade, como na avaliação do projeto. Neste sentido, os resultados obtidos permitem um aumento de eficiência, ou seja, diminuição do tempo de análise das candidaturas devido à utilização de modelos preditivos, e eficácia, onde os fundos são atribuídos apenas a empresas que cumprirão os objetivos contratualizados, do processo de análise das candidaturas.

5.1 Contributos

Perante o problema colocado, o estudo vem contribuir tanto para a academia, como para os profissionais. Por um lado, a RSL efetuada sobre modelos baseados em *accruals* contribui para ajudar os futuros investigadores na área da contabilidade e auditoria, na medida em que é disponibilizado um resumo dos principais artigos publicados em *journals* de mérito reconhecido, e identificados os melhores artigos em quatro dimensões de análise (tipologias, contextos, avaliação dos resultados e qualidade), ou seja, este estudo torna-se um bom ponto de partida para quem pretender estudar a MR.

Também a parte empírica deste estudo revela contributos para o conhecimento científico. Assim, ao incluírem-se quatro modelos de MR baseados em *accruals* evidenciou-se a importância de, ao contrário da tendência, os investigadores contemplarem diversos modelos de MR nos seus estudos, pois cada modelo pode levar a resultados diferentes. Também ao estudar-se a MR no contexto da atribuição de fundos europeus, deu-se a conhecer mais um contexto de aplicação de MBA, quer mais vocacionados para a intensidade, quer para a qualidade desses *accruals*. Por fim, para os investigadores mais focados em técnicas de *data mining/machine learning*, em que se considera, à partida, que não existe uma técnica mais adequada para cada problema ou conjunto de dados, o estudo revela uma forte capacidade preditiva dos modelos de classificação baseados em árvores de decisão e que o algoritmo CHAID, neste contexto específico, levou a modelos de melhor qualidade preditiva. É de salientar, que através da modelação e quando aplicado a fundos, os resultados obtidos com árvores de decisão, foram superiores aqueles obtidos por estudos anteriores através da aplicação de regressão logística, quer de redes neuronais. Ainda relacionado com os modelos de árvores de decisão, o estudo revela que estes modelos se podem aplicar com sucesso nas áreas da contabilidade e auditoria, podendo ser uma melhor alternativa a modelos mais tradicionais ao nível da estatística,

como sejam os modelos de regressão logística, e em que as variáveis preditoras têm que ser independentes.

Para os profissionais, apresenta-se também, numa forma resumida, a problemática e a importância da MR para os diferentes objetivos e contextos, quais os modelos utilizados e avaliação dos resultados obtidos. Quanto aos governantes, é sugerida uma razão para a ineficiência do processo da aplicação de fundos, bem como, são apontadas, aos profissionais do IAPMEI, que avaliam as candidaturas, formas a otimizar o processo (os modelos preditivos). Mais concretamente, é sugerido que estes tenham em consideração a MR, pois os resultados evidenciaram que estas são das variáveis mais importantes, nomeadamente no ano da candidatura e no ano anterior a esta, uma vez que também ficou comprovado que não existem ganhos em obter informação histórica, considerando-se por isso um processo ineficiente. Por outro lado, e a todos os organismos intermédios, e deu-se a conhecer que é possível com boa capacidade preditiva e sem grande esforço de análise prever o sucesso das candidaturas, e, eventualmente, com base nos resultados do modelo não aprovar aquelas candidaturas com forte propensão a não ser elegíveis, mas sendo capaz de explicar aos empresários as razões para a não aprovação. De facto, os modelos baseados em árvores possuem como característica fundamental a explicabilidade.

5.2 Limitações e pistas futuras de investigação

No que concerne às limitações, é importante referir, que a falta de informação afetou o estudo em duas perspetivas diferentes. A primeira, na ausência de informação financeira em relação a algumas empresas, levando a que estas fossem excluídas do estudo. Em segundo, o facto do programa de incentivos ainda estar a decorrer fez com que a não fosse possível obter informação sobre o mérito do projeto, pelo que foi calculado uma *proxy*, isso é, uma aproximação que poderá não espelhar a verdadeira realidade.

Em termos futuros, e para colmatar a falta de rigor da *proxy* do mérito do projeto, recomenda-se replicar o estudo quando o programa, agora em vigor, esteja terminado, utilizando os mesmos algoritmos e incluindo ainda outros, por exemplo as redes neuronais e o C5. Além disso, e de forma a melhorar a qualidade do modelo, sugere-se a inclusão de outras variáveis independentes nos modelos, bem como alargar o estudo a outros programas de incentivos.

Referências Bibliográficas

- ACFE. (2018). *The Report to the Nation on Occupational Fraud and Abuse*, Austin: Association of Certified Fraud Examiners.
- AICEP (2018). *Incentivos Geridos pela AICEP enquanto OI*, Apresentação AICEP, Lisboa: AICEP Portugal Global.
- Atieh, A. & Hussain, S. (2012). Do UK firms manage earnings to meet dividend thresholds? *Accounting and Business Research*, 42 (1): 77-94.
- Balsam, S., Bartov, E., & Marquardt, C. (2002). Accruals management, investor sophistication, and equity valuation: Evidence from 10-Q filings. *Journal of Accounting Research*, 40 (4): 987-1012.
- Baralexis, S. (2004). Creative Accounting in small advancing countries. *Managerial Auditing Journal*, 19 (3): 440-461.
- Beck, A. W. (2018). Opportunistic financial reporting around municipal bond issues. *Review of Accounting Studies*, 23: 785-826.
- Beneish, M. D. (1997). Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy*, 16(3): 271-309.
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5): 24-36.
- Beneish, M. D. (2001). Earnings management: A perspective. *Managerial Finance*, 27(12): 3-17.
- Beneish, M. D. (2005). *Earnings Quality and Future Returns: The Relation between Accruals and the Probability of Earnings Manipulation*. Indiana.
- Beneish, M.D., Lee, C.M.C., & Nichols, D.C. (2013). Earnings manipulation and expected returns. *Financial Analysts Journal*, 69(2): 57-82.
- Boone, J. P., Khurana, I. K. & Raman, K. K. (2011). Litigation Risk and Abnormal Accruals. *Auditing: A Journal Of Practice & Theory*, 30 (2): 231-256.
- Boynton, C. E., Dobbins, P. S., & Plesko, G. A. (1992). Earnings management and the corporate alternative minimum tax. *Journal of Accounting Research*, 30: 131-153.
- Burgstahler, D., & Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24: 99-126.
- Breiman, L., Friedman, J. H., Olshen, R. J., & Stone, C. J. (1984). *Classification and Regression Trees*. Belmont, California. Wadsworth International Group.
- Caldas, R., Mundt, M., Potthast W., Neto, F. ,& Markert, B. (2017). A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. *Gait & Posture*, 57: 204-210.
- Callao, S., & Jarne, J. I. (2010). Have IFRS affected earnings management in the European Union? *Accounting in Europe*, 7(2): 159-189.
- Carmo, C., Moreira, J., & Miranda, M. (2010) A Qualidade dos Accruals e o Custo do Financiamento nas Empresas Portuguesas uma Análise por Grupos de Dimensão. In Ata do 11º Congresso Internacional AECA.
- Carmona Ibanez, P. & Momparler Pechuan, A. (2011). Nonaudit services provided by incumbent auditors and earnings management: Evidence of auditor independence from an EU country. *Spanish Journal of Finance and Accounting*, 40 (152): 587-612.
- Carver, B.T., Hollingsworth, C. W. & Stanley, J. D. (2011). Recent Auditor Downgrade Activity and Changes in Clients' Discretionary Accruals. *Auditing: A Journal Of Practice & Theory*, 30 (3): 33-58.
- Chang, C. & Lin, H. (2018). Does there prevail momentum in earnings management for seasoned equity offering firms? *International Review of Economics & Finance*, 55: 111-129.
- Chen, C., Lin, C., & Lin, Y. (2008). Audit partner tenure, audit firm tenure, and discretionary accruals: Does long auditor tenure impair earnings quality? *Contemporary Accounting Research*, 25 (2): 415-445.
- Chi, H.& Chin, C. (2011). Firm versus Partner Measures of Auditor Industry Expertise and Effects on Auditor Quality. *Auditing: A Journal Of Practice & Theory*, 30 (2): 201-229.

- Choi, J., Kim, J., & Zang, Y. (2010). Do Abnormally High Audit Fees Impair Audit Quality? *Auditing: A Journal Of Practice & Theory*, 29 (2): 115-140.
- Constantino, J. (2020). *Gestão de Resultados como Preditor de Insolvências: Evidência nas Empresas Portuguesas*, Dissertação de Mestrado, ISCTE-IUL, Lisboa.
- Davidson, S., Stickney, C. & Weil, R. (1987). *Accounting: The Language of Business*. Thomas Horton and Daughters, Sun Lakes, Arizona.
- De Fuentes, C., Illueca, M. & Consuelo Pucheta-Martinez, M. (2015). External investigations and disciplinary sanctions against auditors: the impact on audit quality. *SERIEs*, 6: 313-347.
- DeAngelo, L. (1986). Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders, *The Accounting Review*, 61: 400-420
- Dechow, P. & Sloan, R. (1991). Executive incentives and the horizon problem: An empirical investigation. *Journal of Accounting and Economics*, 14: 51-89.
- Dechow, P., Sloan, R. & Sweeney, A. (1995). Detecting earnings management. *The Accounting Review*, 70: 193-225.
- Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the sec. *Contemporary Accounting Research*, 13(1): 1-36.
- Dechow, P. M., & Dichev, I. D. (2002). The Quality of accruals and Earnings: The Role of Accrual Estimation Errors. *Accounting Review*, 77(4): 35.
- DeFond, M., & Jiambalvo, J. (1991). Incidence and Circumstances of Accounting Errors. *The Accounting Review*, 66: 643-655.
- DeGeorge, F., Patel, J., & Zeckhauser, R. (1999). Earnings management to exceed thresholds. *Journal of Business* 72: 1-33.
- Delen D., Kuzey C. & Uyar A. (2013). Measuring firm performance using financial ratios: a decision tree approach. *Expert Systems with Applications*, 40(10): 3970-3983.
- Dias, P. (2015). *As diferenças entre o resultado contabilístico e o fiscal e a gestão dos resultados: Evidência empírica de empresas privadas portuguesas*. Dissertação de Doutoramento, ISCTE-IUL, Lisboa.
- ECA. (2017). *EU audit in brief. (2016)*. Publications Office of the European Union, European Court of Auditors.
- Jensen, M. & Meckling, W. (1976). Theory of the firm: Managerial behaviour, agency costs and ownership structure. *Journal of Financial Economics*, 3(4): 305-360.
- Fernandes, S. (2011). *O uso de data mining na previsão do desempenho das empresas portuguesas: Estudo de casos no âmbito do SIME – Sistema de Incentivos à Modernização Empresarial*. Dissertação de Mestrado, ISCTE-IUL, Lisboa.
- Fernandes, S., & Laureano, R. (2019). *Manipulação de resultados e fraude nas candidaturas a fundos europeus: estudo em empresas portuguesas*. In Ata do 20º Congresso Internacional AECA, Málaga.
- Fisher, T. C. G., Gavius, I. & Martel, J. (2019). Earnings Management in Chapter 11 Bankruptcy. *ABACUS*, 55 (2): 273-305.
- Geiger, M. A. & North, D. S. (2006). Does hiring a new CFO change things? An investigation of changes in discretionary accruals. *Accounting Review*, 81 (4): 781-809.
- Gill-De-Albornoz, B. & Illueca, M. (2005). Earnings management under price regulation: Empirical evidence from the Spanish electricity industry. *Energy Economics*, 27: 279 – 304.
- Guenther, D. (1994). Earnings Management in Response to Corporate Tax Rate Changes: Evidence from the 1986 Tax Reform Act. *The Accounting Review*, 69 (1): 230-243.
- Habib, A. & Bhuiyan, M. (2016). Problem directors on the audit committee and financial reporting quality. *Accounting and Business Research*, 46 (2): 121-144.
- Healy, P. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, 7 (1-3): 85-107.
- Healy, P. M., & Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, 13(4): 365-383.
- INE. (2007). *Classificação Portuguesa das Atividades Económicas Rev.3*: 39-42.

- Im, C. & Nam, G. (2019). Does Ethical Behavior of Management Influence Financial Reporting Quality? *Sustainability*, 11 (20): 5765-5781.
- Ittonen, K., Vahamaa, E. & Vahamaa, S. (2013). Female Auditors and Accruals Quality. *Accounting Horizons*, 27 (2): 205-228.
- Jacinto, S. M. (2019). *Determinantes da auditoria voluntária nas PME portuguesas*. Dissertação de Mestrado, ISCTE-IUL, Lisboa.
- Jackson, A. B. (2017). Discretionary Accruals: Earnings Management ... or Not? *ABACUS*, 54 (2): 136-153.
- Jensen, M. & Meckling, W. H. (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3 (4): 305-360.
- Jha, A. (2019). Financial Reports and Social Capital. *Journal of Business Ethic*, 155: 567-596.
- Jones, J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29 (Autumn): 193-228.
- Jones, K.L., Krishnan, G. V., & Melendrez, K. D. (2008). Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical analysis. *Contemporary Accounting Research*, 25 (2): 499-531.
- Kalyta, P. (2009). Accounting Discretion, Horizon Problem, and CEO Retirement Benefits. *Accounting Review*, 84 (5): 1553-1573.
- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29(2): 119-127.
- Kaszniak, R. (1999). On the association between voluntary disclosure and earnings management, *Journal of Accounting Research*, 37(1): 57-82.
- Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of accounting and economics*, 39(1): 163-197.
- Kurt, A.C. (2018). How Do Financial Constraints Relate to Financial Reporting Quality? Evidence from Seasoned Equity Offerings. *European Accounting Review*, 27 (3): 527-557.
- Larcker, D. F., & Richardson, S. A. (2004). Fees paid to audit firms, accrual choices, and corporate governance. *Journal of Accounting Research*, 42(3): 625-658.
- Larcker, D., Richardson, S., & Tuna, I. (2007). Corporate Governance, Accounting Outcomes, and Organizational Performance. *The Accounting Review* 82 (4): 963-1008.
- Laureano, R. (2020). *Testes de Hipóteses e Regressão: O Meu Manual de Consulta Rápida*, Lisboa: Edições Sílabo.
- Lisboa, I. (2016). Impact of financial crisis and family control on earning management of Portuguese listed firms. *European Journal of Family Business*, 6(2): 118-131.
- Lin, W. & Liao, T. (2015). Exchange listing type and firm financial reporting behavior. *International Review of Economics & Finance*, 38: 234-249.
- Lin, W. & Liao, T. (2018). Managerial reporting behavior around exchange switching: Consideration of current and future performance. *International Review of Economics & Finance*, 56: 218-237.
- Louis, H. & Robinson, D. (2005). Do managers credibly use accruals to signal private information? Evidence from the pricing of discretionary accruals around stock splits. *Journal of Accounting & Economics*, 39: 361-380.
- Lourenço, I., Morais, A. I., & Lopes, A.I. (2015). *Fundamentos de contabilidade financeira: teoria e casos*. Lisboa: Edições Sílabo.
- Major, M. J. (2009). Reflexão sobre a investigação em Contabilidade de Gestão. *Revista de Gestão dos Países de Língua Portuguesa*, 8(1): 43-50.
- Major, M. J. (2017). O positivismo e a pesquisa 'alternativa' em Contabilidade. *Revista Contabilidade & Finanças*, 28 (74): 173-178.
- Marques, M. & Rodrigues, L. (2009). A Problemática do Earnings Management e suas Implicações. *Jornal de Contabilidade da APOTEC*: 325-339.
- McNichols, M., & G. P. Wilson (1988). Evidence of Earnings Management from the Provision for Bad Debts. *Journal of Accounting Research*, 26: 1-31.

- McNichols, M. F. (2000). Research design issues in earnings management studies. *Journal of Accounting and Public Policy*, 19: 313-345.
- McNichols, M. F. (2002). Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *The accounting review*, 77: 61- 69.
- Michalewicz, Z., Schmidt, M., Michalewicz, M., & Chiriack, C. (2006). *Adaptive Business Intelligence* (1 ed.). New York.
- Mingers, J. (1989). An empirical comparison of selection measures for decision-tree induction. *Machine Learning*, 3: 319-342.
- Moreira, J. A. (2008). A Manipulação dos Resultados nas Empresas: um contributo para o estudo do caso Português, *Jornal de Contabilidade da APOTEC*: 144-153.
- Mulford, C. W., & Comiskey, E. E. (2011). *The financial numbers game: detecting creative accounting practices*. John Wiley & Sons.
- Myers, J. N., Myers, L. A. & Omer, T. C. (2003). Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *Accounting Review*, 78 (3): 779-799.
- OLAF. (2018). *The OLAF report 2018*. Luxemburgo: Publications Office of the European Union.
- OLAF. (2015). *The role of Member States' auditors in fraud prevention and detection*. European Union. Directorate D — Policy, European Anti-Fraud Office.
- Penman, S. H. (2012). *Financial Statement Analysis and Security Valuation*, Irwin Professional Pub.
- Peasnell, K. V., Pope, R.F., & Young, S. (2005). Board monitoring and earnings management: Do outside directors influence abnormal accruals? *Journal of Business Finance & Accounting*, 32 (7-8): 1311-1346.
- Pereira, J. M., Basto, M., Díaz Gómez, F., & Barbas Albuquerque, E. (2010). *Los Modelos de Predicción del Fracaso Empresarial*. Propuesta de un Ranking, XIV encuentro AECA, Coimbra, Portugal.
- Perotti, P. & Windisch, D. (2017). Managerial Discretion in Accruals and Informational Efficiency. *Journal of Business Finance & Accounting*, 44 (3-4): 375-416.
- Pestana, H. M. & Gageiro, J. N. (2014). *Análise categórica, árvores de decisão e análise de conteúdo em ciências sociais e da saúde com o SPSS*. (6ª ed.). Lisboa: Edições Sílabo.
- Pires, L. M. (2017). 30 anos de Fundos Estruturais (1986-2015). *Relações Internacionais (R:I)*, (53): 19-38.
- Price, Richard A., Sharp, N. Y., & Wood, D. A. (2011). Detecting and Predicting Accounting Irregularities: A Comparison of Commercial and Academic Risk Measures. *Accounting Horizons*, 25 (4): 755-780.
- Qi, B., Yang, R. & Tian, G. (2017). Do social ties between individual auditors and client CEOs/CFOs matter to audit quality? *Asia-Pacific Journal of Accounting & Economics*, 24 (3-4): 440-463.
- Ravenda, D., Valencia-Silva, M. M., Maria Argiles-Bosch, J. & Garcia-Blandon, J. (2018). Accrual management as an indication of money laundering through legally registered Mafia firms in Italy. *Accounting, Auditing & Accountability Journal*, 31 (1): 286-317.
- Rokach, L., & Maimon, O. (2008). *Data Mining with Decision Trees: Theory and Applications*. Series in Machine Perception and Artificial Intelligence, Vol.69. Hackensack, New Jersey: World Scientific Publishing Co. Pte. Ltd
- Romanus, R. N., Maher, J. J. & Fleming, D. M. (2008). Auditor Industry Specialization, Auditor Changes, and Accounting Restatements. *Accounting Horizons*, 22 (4): 389-413.
- Rosner, R. L. (2003). Earnings manipulation in failing firms. *Contemporary Accounting Research*, 20 (2): 361-408.
- Ronen, J., & Yaari. V. (2007). *Earnings Management: Emerging Insights in Theory, Practice, and Research*, Springer Series in Accounting Scholarship
- Salehi, M., Tarighi, H. & Sahebkar, H. (2018). The impact of auditor conservatism on accruals and going concern opinion: Iranian angle. *International Journal of Islamic And Middle Eastern Finance and Management*, 11 (4): 650-666.
- Schipper, K. (1989). Commentary on earnings management. *Accounting horizons*, 3(4), 91-102.
- Simpson, A. (2013). Does Investor Sentiment Affect Earnings Management? *Journal of Business Finance & Accounting*, 40 (7-8): 869-900.

- Sousa, J. S. & Baptista, C. S. (2011). *Como fazer investigação, dissertações, teses e relatórios segundo Bolonha* (2ª ed.). Lisboa: PACTOR - Edições de ciências sociais e políticas contemporâneas.
- Stec, M. & Grzebyk, M. (2018). The Implementation of the Strategy Europe 2020 Objectives in European Union Countries: The Concept Analysis and Statistical Evaluation. *International Journal of Methodology - Quality & Quantity*, 52(1): 119-133.
- Sweeney, A. P. (1994). Debt Covenant Violations and Managers' Accounting Responses. *Journal of Accounting and Economics*, 17: 281-308.
- Talab, H., Flayyih, H., & Ali, S. (2017). Role of Beneish M-score Model in Detecting of Earnings Management Practices: Empirical Study in Listed Banks of Iraqi Stock Exchange. *International Journal of Applied Business and Economic Research*, 15 (23): 287-302.
- Teoh, S. H., Welch, I., & Wong, T. J. (1998). Earnings management and the underperformance of seasoned equity offerings. *Journal Of Financial Economics*, 50: 63-99.
- Vilelas, J. (2017). *Investigação: O Processo de Construção do Conhecimento* (2ª ed.). Lisboa: Edições Sílabo
- Watts, R. L., & Zimmerman, J. L. (1986). *Positive accounting theory*. Prentice-Hall Inc.
- West, J. & Bhattacharya, B. (2016). Some Experimental Issues in Financial Fraud Mining. *Procedia Computer Science*, 80: 1734-1744.
- Wong, R. M. K., Lo, Ag. W. Y. & Firth, M. (2015). Managing Discretionary Accruals and Book-Tax Differences in Anticipation of Tax Rate Increases: Evidence from China. *Journal of International Financial Management & Accounting*, 26 (2): 188-222.