



UNIVERSIDADE CATÓLICA PORTUGUESA

Modelo de notação de risco para famílias portuguesas

Trabalho Final na modalidade de Relatório de Estágio
apresentado à Universidade Católica Portuguesa
para obtenção do grau de mestre em Banca e Seguros

por

Catarina Andreia Nogueira Henriques

sob orientação de

Professor Doutor Manuel Ricardo Fontes da Cunha
Professor Doutor Paulo Alexandre Pimenta Alves

Faculdade de Economia e Gestão
Setembro, 2014

Agradecimentos

Aos meus Pais e irmão por me apoiarem e aconselharem nas minhas decisões e ambições.

Ao professor Ricardo Cunha e ao professor Paulo Alves pela orientação nesta dissertação.

Ao Banco de Portugal e ao Departamento de Estatística pelo estágio proporcionado.

Às coordenadoras do estágio no Banco de Portugal, Doutora Ana Pinto coordenadora do Núcleo de Análise de Balanços e à Doutora Paula Casimiro coordenadora da Área de Estatísticas Monetárias e Financeiras e de Centralização das Responsabilidades de Crédito.

Às Doutoras Sónia Costa e Luísa Farinha, do Departamento de Estudos Económicos do Banco de Portugal, pelos esclarecimentos relativos ao Inquérito à Situação Financeira das Famílias.

Ao Hélder, colega de estágio, que esteve presente nesta etapa e que me aconselhou em muitos momentos e a todos os meus amigos.

Lista de abreviaturas

a.C. – antes de Cristo	HFCN – <i>Household Finance and Consumption Network</i>
AD – Análise Discriminante	IMF – <i>International Monetary Fund</i>
AIC – <i>Akaike information criterion</i>	INE – Instituto Nacional de Estatística
CE – Conselho Europeu	ISFF – Inquérito à Situação Financeira da Famílias
CRC – Central de Responsabilidades de Crédito	LGD – <i>Loss Given Default</i>
CPR – Companhia Portuguesa de Rating, SA	OCC – <i>Office of the Comptroller of the Currency</i>
BCE – Banco Central Europeu	PD – <i>Probability of default</i>
BdP – Banco de Portugal	PIB – Produto Interno Bruto
DB – <i>Dun & Bradstreet</i>	RL – Regressão Logística
EAD – <i>Exposure At Default</i>	RN – Redes Neurais
EL – <i>Expected Loss</i>	ROC – <i>Receiver Operating Characteristic</i>
ESMA – Autoridade Europeia dos Valores Mobiliários e dos Mercados	RP – Regressão Probabilística
EUA – Estados Unidos da América	RPA – Residência Principal do Agregado
REF – Relatório de Estabilidade Financeira	RT – <i>Regression Tree</i>
FICO – <i>Fair Isaac Company</i>	SaeR – Sociedade de Avaliação de Empresas e Risco
HFCN – <i>Household Finance and Consumption Survey</i>	UE – União Europeia
	VIF – Fator de Inflação das Variâncias

Resumo

O risco de crédito e a sua mensuração são o principal foco de uma instituição de crédito, facto em muito despoletado pela crise financeira atual, que trouxe consigo restrições de crédito, fruto da menor liquidez dos mercados financeiros mundiais e acrescido do elevado número de incumprimentos verificados, tanto pelas pessoas coletivas como singulares. Este elevado incumprimento advém, entre muitos outros fatores, da má medição de risco de crédito que se traduz no financiamento de devedores com elevadas probabilidades de incumprir as suas obrigações.

Este trabalho incide na medição da probabilidade de incumprimento de pessoas singulares. Foi assim estimado um modelo econométrico para determinar uma notação de risco de crédito, partindo de dados do Inquérito à Situação Financeira das Famílias realizado em 2010, em Portugal. Foi possível não só determinar a probabilidade de incumprimento de uma família, como também determinar a exposição a incumprimento e a perda em caso de incumprimento.

O modelo proposto teve como base uma regressão probabilística objeto de calibração, com base na junção de variáveis explicativas já estudadas por outros autores. Para confirmar a solidez do modelo realizou-se uma análise estatística, de precisão, ajuste e respetiva validação do mesmo.

Concluiu-se que das variáveis explicativas incluídas no modelo, o número de crianças, a recusa de crédito e a deterioração das condições contribuem de forma significativa e positiva para a probabilidade de incumprir. Por outro lado, a propriedade de habitação contribui para a diminuição da probabilidade de incumprimento de um agregado familiar.

Palavras-chave: crédito, risco de crédito, modelo de notação de risco de crédito e agências de notação de risco de crédito

Abstract

Credit risk and its measurement are the primary focus of a credit institution, especially since the start of the current financial crisis, which has brought credit constraints to the illiquidity of markets and a large number of defaults observed for firms and individuals. This failure arises from poor measurement of credit risk which translates in financing borrowers with high probabilities of default.

This dissertation focuses on the measurement of the household probability of default. We estimate an econometric model to determine a credit rating, based on the data from the Household Finance and Consumption Survey held in 2010, in Portugal. It was possible not only to determine the household probability of default, but also to determine the exposure to default and the loss given default.

The proposed model was based on a probabilistic regression including independent variables studied by other authors. To confirm the robustness of the model a thorough statistical analysis, and its was developed adjustment, accuracy and validation were tested.

It was concluded that the independent variables included in the model, the number of children, credit declined and the deteriorating credit conditions contributed significantly and positively to the probability of default. On the other hand, homeownership contributes to the decrease of the household probability of default.

Keywords: credit, credit risk, credit scoring model and credit rating agencies

Índice

Agradecimentos	iii
Lista de abreviaturas	v
Resumo	vii
Abstract	viii
1 Introdução.....	1
2 Risco de Crédito e Notação	7
2.1 Conceitos e história	7
2.1.1 Risco de crédito	9
2.1.2 Crédito a particulares em Portugal	12
2.2 Notação	15
2.2.1 Notação de risco	16
2.3 Modelos de notação de risco.....	22
2.3.1 Modelos estatísticos.....	24
2.3.2 Etapas dos modelos de notação de risco	30
2.3.3 Breve nota aos estudos realizados em Portugal	32
2.4 Agências de notação de risco.....	36
2.4.1 Tipos de notação das agências de notação de risco	36
2.4.2 Tipos de agências de notação	37
2.4.3 Agências de notação em Portugal	41
3 Modelo de notação de risco.....	45
3.1 Caracterização da amostra.....	45
3.1.1 Caracterização e principais resultados do ISFF 2010.....	46

3.2 Seleção das variáveis	47
3.2.1 Variável dependente.....	48
3.2.2 Variáveis explicativas.....	48
3.2.3 Resumo das variáveis.....	54
3.3 O Modelo.....	56
3.4 Resultado Empírico.....	58
3.4.1 Testes empíricos e ajustamento ao modelo.....	58
3.4.2 Construção do modelo.....	59
3.4.3 O Modelo Final.....	66
3.4.4 Avaliação do modelo proposto.....	67
3.5 Classes de risco.....	71
3.6 Perda esperada.....	74
4 Conclusões.....	76
5 Limitações do estudo.....	79
6 Pesquisa futura e sugestões.....	80
Referências bibliográficas.....	82
Legislação.....	91
Anexos.....	92

Índice de Tabelas

Tabela 1 – Vantagens genéricas dos modelos estatísticos de notação de risco...	24
Tabela 2 – Vantagens comparativas dos vários modelos estatísticos de notação de risco	28
Tabela 3 – Limitações dos vários modelos estatísticos de notação de risco.....	29
Tabela 4 – Componentes e pesos dados às variáveis do FICO score <i>versus</i> VantageScore.....	40
Tabela 5 – Caracterização das variáveis selecionadas.....	54
Tabela 6 – Estatísticas descritivas das variáveis selecionadas	55
Tabela 7 – Regressão probabilística para probabilidade de incumprimento das famílias (Modelo 1).....	60
Tabela 8 – Regressão probabilística para probabilidade de incumprimento das famílias (Modelos 2, 3 e 4)	62
Tabela 9 – Regressão probabilística para probabilidade de incumprimento das famílias (Modelo 5).....	64
Tabela 10 – Passos do modelo proposto.....	66
Tabela 11 – Teste de Hosmer-Lemeshow	67
Tabela 12 – Matriz de classificação do modelo proposto	68
Tabela 13 – Teste de Breusch-Pagan	69
Tabela 14 – Matriz de classificação resultado da validação do modelo	70
Tabela 15 – Intervalos de classificação para probabilidade de incumprimento.	72

Índice de Figuras

Figura 1 – Cálculo da perda esperada	10
Figura 2 – Tipos de modelos usados na medição de risco.....	23
Figura 3 – Cronograma das principais agências de <i>rating</i>	38

Índice de Gráficos

Gráfico 1 – Crédito a particulares em Portugal (variação anual 2006-2013)	12
Gráfico 2 – Rácio de crédito vencido e crédito a particulares (2009-2014)	13
Gráfico 3 – Dívida total em relação ao rendimento das famílias	14
Gráfico 4 – Distribuição da variável dependente	48
Gráfico 5 – Curva ROC	71
Gráfico 6 – Distribuição da amostra pelas classes de risco	73

1 Introdução

O sistema financeiro é essencial para o funcionamento da sociedade como a conhecemos, com implicações diretas na riqueza, no funcionamento das empresas e na eficiência económica (Mishkin 2004). Conforme enunciado pelo mesmo autor, o sistema financeiro é complexo em termos de estrutura e função, tendo como finalidade a alocação de recursos financeiros, entre os diversos agentes da economia, de onde se destacam, com especial importância para este estudo, os bancos, as seguradoras, os fundos de investimento, as ações e mercados de títulos, tal como as entidades destacadas para a sua regulação e supervisão.

A evolução do sistema financeiro assentou na crescente necessidade de financiamento, aliada à maior distância entre investidores e os que necessitavam de financiamento, deixando de estar confinado a um ponto de encontro físico entre os diversos agentes, para se tornar também um encontro de vontades expressas eletronicamente.

Contudo estas transações entre financiado e financiador não são isentas de risco, risco este que se prende com a possibilidade de uma das partes não cumprir os compromissos assumidos inicialmente.

A crescente procura de financiamento arrastou consigo a preocupação com o risco de crédito e a sua medição. A avaliação do risco de crédito tornou-se, então, foco de atenção crescente, em muito justificada pela importância dada aos mercados financeiros, despoletada pela crise financeira de 2007, que teve a sua origem nos EUA e que depressa contagiou outros mercados, especialmente o Europeu. Esta atenção foi simultaneamente dirigida a todos os agentes do mercado, desde os que os influenciam (como as instituições financeiras, investidores, *etc.*) aos que têm a missão de o regular.

Entre os agentes que influenciam o mercado inserem-se as agências de notação de risco, que surgiram há pouco mais de 60 anos, com o intuito,

conforme definição do Regulamento (CE) n.º 1060/2009, de emitir “um parecer relativo à qualidade de crédito de uma entidade, de uma obrigação de dívida ou obrigação financeira, de títulos de dívida, de ações preferenciais ou outros instrumentos financeiros, ou do emitente de tais obrigações de dívida ou obrigações financeiras, títulos de dívida, ações preferenciais ou outros instrumentos financeiros, emitido através de um sistema de classificação estabelecido e definido com diferentes categorias de notação”.

O *International Monetary Fund* (IMF) (2010) acrescenta ainda que são dois os principais serviços prestados pelas agências de notação: um serviço de informação ao fornecer uma avaliação independente da capacidade dos emitentes cumprirem as suas obrigações de dívida; e um serviço de monitorização, ao sugerirem, aos emitentes, a tomada das ações corretivas capazes de melhorar ou manter a notação.

A notação de risco, presentemente, é despoletada por interesse do devedor (Butler e Cornaggia, 2012), sendo uma forma de sinalizar aos seus potenciais investidores a sua condição geral de solvabilidade, não sendo todavia isenta de erros (Investe, 2009), nem deve ser tida como recomendação de investimento.

De acordo com a Autoridade Europeia dos Valores Mobiliários e dos Mercados (ESMA) (ESMA, 2014a e 2014b), neste momento encontram-se registadas 23 agências de notação de risco e duas agências certificadas, a atuar no espaço da União Europeia (UE), onde se destaca a primazia de 3 grandes agências americanas – Moody’s, Standard & Poor’s e Fitch Rating – que se dedicam ao desenvolvimento de notação de risco de países, empresas e produtos financeiros. Em Portugal existe apenas uma agência de notação de risco, a ARC Ratings, que é dirigida a empresas e instituições financeiras.¹

¹ A diferença entre agências de notação de risco certificadas e registadas reside no local de estabelecimento das agências e de quem faz a sua supervisão, de tal forma que agências com estabelecimento fora comunidade europeia podem ser certificadas desde que: a) esteja autorizada ou registada e objeto de supervisão no país terceiro em causa; b) a Comissão reconheça a equivalência do enquadramento legal e de supervisão do país terceiro em causa; c)

A análise de risco não é foco somente das agências de notação de risco, pois também as instituições de crédito avaliam o risco dos seus clientes. Contudo é de ressaltar que nestas últimas o *core* da sua atividade é a intermediação financeira, seja esta de prazos ou montantes. Assim se compreende que as instituições não dispensem as notações emitidas pelas agências de notação de risco, acreditando que estas são isentas e prospectivas de cumprimento dos compromissos futuros.

Maioritariamente as agências de notação de risco concentram-se no desenvolvimento e divulgação de notações dirigidas a empresas ou países, não sendo tão comum a avaliação de risco de particulares. A análise de risco de particulares recai então unicamente sobre as instituições de crédito, sempre que têm de conceder financiamento a um particular, cingindo-se à informação que o cliente disponibiliza, ao histórico enquanto cliente e ao conhecimento que anos de atividade proporcionam nesta matéria.

Admite-se que, no caso específico dos particulares, seja difícil o acesso a todas as variáveis necessárias para a medição do risco de financiamento (Medina e Selva, 2012), mesmo que, segundo Abdou e Pointon (2011), não exista um número ideal de variáveis para analisar risco de crédito, seja devido ao modelo de notação de risco usado, seja por questões de proteção de informação do particular, que em Portugal está sob a alçada Comissão Nacional de Proteção de Dados. No entanto, esta organização começa a dar visibilidade a informações que até então eram restritas, desde listagem de devedores à Autoridade Tributária, à Segurança Social e até a serviços de telecomunicações.

existam acordos de cooperação com as autoridades competentes dos países terceiros cujos enquadramentos legais e de supervisão tenham sido considerados equivalentes; e ainda que d) as notações de risco emitidas e as suas atividades de notação de risco não tenham importância sistémica para a estabilidade financeira ou a integridade dos mercados financeiros de um ou mais Estados-Membros (conforme artigo nº 5 do Regulamento (CE) nº 1060/2009 do Parlamento Europeu e do Conselho). Todas as restantes agências, em território Europeu e sujeitas à supervisão da ESMA, são consideradas registadas.

Nesta matéria os EUA foram os primeiros a desenvolver um modelo de avaliação e previsão de risco também para particulares, desenvolvido pela *Fair Isaac Company* (FICO) e denominado por FICO score (FICO, 2014a). Esta metodologia rapidamente se tornou uma importante ferramenta de gestão de risco (FICO, 2014b), quer para as instituições financeiras, quer para os próprios particulares, que viram na sua notação de risco um veículo facilitador da relação com as instituições financeiras (Lyons *et al.*, 2007). O mérito desta notação dirigida a particulares é justificado, pela própria *Fair Isaac Company*, pela promoção da concessão de empréstimos mais justos, pela facilidade de compreensão pelos reguladores, pelo cumprimento dos padrões definidos, da mesma forma que se adapta a mudanças regulatórias e que é mais transparente no que toca às variáveis que estão subjacentes.

O presente Trabalho Final de Mestrado tem por objetivo primordial a elaboração de um modelo de notação de risco para particulares adaptado ao mercado português, que permita uma correta medição do risco de particulares e que seja um bom auxiliar das instituições financeiras, no que concerne à concessão de novos financiamentos para particulares.² A tomada de uma decisão de crédito, tal como a respetiva determinação de uma notação de risco, é tida como um “*Puzzle Sudoku*” segundo a *National Association for Court Management* (2011)³, já que são muitas as peças a considerar, o que na

² Ao longo desta análise será usado o termo notação de risco comumente conhecido como *scoring* ou *rating*. O termo *rating* por regra é usado para notações de risco que incidem sobre estados, países ou empresas cotadas em que são usados modelos de análise mais complexos e incluem variáveis características deste tipo de sujeitos. Por outro lado, o termo *scoring* é usado para notações que incidem sobre pequenas empresas e particulares.

Na prática *scoring* e *rating* têm o mesmo objetivo que dar um parecer quanto à qualidade de crédito, a diferença está na escala usada, enquanto o *rating* é comumente refletido numa escala de letras e símbolos (em alguns casos também inclui números), a escala de *scoring* é apenas numérica.

³ *National Association for Court Management* (NACM) é uma associação constituída por membros de vários países desde os EUA, Canadá, Austrália, *etc.*

inexistência de algumas poderá implicará erros na análise e na notação atribuída.

Muitos são os autores que emergem com análises subjetivas (por exemplo Santos, 2003) sobre o tipo de variáveis a ter em consideração na análise de risco de crédito, realizando desde análises puramente estatísticas até análises teóricas que refletem a experiência do analista e o seu conhecimento, que nem sempre contêm todas as características e qualidades de determinado candidato. Depois de selecionadas e recolhidas as variáveis, sucede-se a análise das mesmas. É nesta etapa que o objetivo deste estudo tem um grande peso na escolha da metodologia a aplicar.

Muitos são também os modelos usados quando se estuda esta temática, recaindo, maioritariamente, em análises econométricas, que podem passar pelo uso de modelos de regressão linear, análise discriminante, regressão logística, árvores de decisão, redes neuronais e programação linear (Anderson, 2007).

O modelo proposto terá por base os dados recolhidos no Inquérito à Situação Financeira da Famílias realizado em 2010, desenvolvendo-se uma regressão probabilística, que permitiu determinar quais as variáveis que melhor explicam a probabilidade de incumprimento e qual o seu impacto.

Inicialmente, no Capítulo 2, será feito um enquadramento teórico, sendo realizada uma revisão da literatura sobre os conceitos subjacentes ao tema deste estudo, desde a noção de crédito, risco de crédito, notação de risco, agências de notação, tal como os modelos usados no desenvolvimento de notações de risco e análise de estudos que usaram bases de dados similares à usada no modelo proposto e alusão ainda aos tipos de notação e agências de notação de risco.

O Capítulo 3 destina-se à apresentação da metodologia em que assentará toda a análise e estimação do modelo proposto, sendo apresentado o modelo proposto tal como as linhas de orientação essenciais para a sua condução. A descrição, do mesmo, passa pela apresentação da base de dados, discriminação das variáveis selecionadas, discussão dos resultados obtidos e respetivas

conclusões. Neste capítulo são apresentados os testes empíricos e de ajustamento, como foi construído o modelo e a respetiva avaliação. Ainda, no Capítulo 3, são definidas classes de risco e determinado o valor de perda esperada.

O Capítulo 4 destina-se à exposição das conclusões encontradas ao longo do desenvolvimento deste estudo, no Capítulo 5 tem lugar a apresentação das limitações encontradas, e por fim no Capítulo 6 salientam-se pontos importantes para pesquisa futura e sugestões que enriqueceriam este estudo.

2 Risco de Crédito e Notação

Este Capítulo apresenta os conceitos e os modelos que suportam este estudo, passando pelo conceito de crédito, risco de crédito, notação e notação de risco, agências de notação, os modelos de notação comumente desenvolvidos em matéria de análise de risco de crédito de particulares, e ainda um breve enquadramento do crédito e das agências em Portugal.

2.1 Conceitos e história

Recuando ao início da história económica, o conceito de crédito surge em 2000 a.C. na Babilónia, Assíria e Egito, embora registos que datam de 3000 a.C. revelem o uso do crédito, com base em cereais e metais (Homer e Sylla, 2005 e Anderson, 2007).

A palavra crédito deriva da palavra *credo* (do latim) que significa confiar. Para confiar é preciso conhecer e para conhecer é necessário tempo e informação (Paiva, 1997). Assim se compreende que uma relação de crédito tenha subjacente a confiança no devedor. Deste modo, o crédito nada mais é do que uma construção humana (Anderson, 2007). Segundo Schreiner (2000), crédito consiste na disposição de alguém ceder, temporariamente, parte do seu património a um terceiro, tendo a expectativa de que tal valor volte à sua posse integralmente, no término do prazo determinado. Na prática crédito refere-se, segundo Santos (2003), à troca de um valor presente por um valor no futuro, dependente sempre do “fator risco”.

Genericamente distinguem-se dois tipos de crédito, com base em quem o concede, o que é concedido por empresas não financeiras e o realizado por instituições financeiras.

O crédito é um dos componentes imprescindíveis para o crescimento de qualquer economia (Bernanke *et al.*, 1996) e, portanto, merece atenção

redobrada, já que tem associada a assunção de possíveis riscos para quem o concede, promove o consumo e o investimento (Stiglitz e Weiss, 1981 e Pereira, 2011). Ao promover o consumo o crédito potência uma maior necessidade de financiamento, esta maior necessidade de consumir resultou claramente do desenvolvimento e modernização da sociedade. O crédito melhorou o conforto e a qualidade de vida, ao permitir satisfazer as necessidades pessoais e sociais (Moreira, 2011).

Juntamente com o conceito de crédito é pertinente esclarecer vários outros conceitos, estreitamente ligados à concessão de crédito, como assimetria de informação, seleção adversa, *creditworthiness*, *moral hazard*, e risco de crédito. Será dada uma breve alusão aos referidos conceitos com uma maior atenção para o risco de crédito, a que é dedicado o subcapítulo seguinte.

O conceito de assimetria de informação deriva de o devedor ser à partida quem mais tem informação sobre si, informação que o concedente não detém (nem é suposto deter) na sua plenitude, assim se compreende que os maus devedores afastem os bons devedores do mercado (Akerlof, 1970), ao levarem os concedentes a serem mais cautelosos e a cobrar mais pela concessão de um crédito. Por outro lado, a própria assimetria de informação pode ser também gerada pelas agências de notação e instituições financeiras sempre que retêm informação da concorrência (Pereira, 2009). Desta forma, segundo Mishkin (2004), a assimetria de informação consiste na insuficiência de conhecimento da outra parte, propiciando a existência de seleção adversa e de *moral hazard*. Segundo Stiglitz e Weiss (1981), o racionamento do crédito é consequência da assimetria de informação das instituições financeiras face aos seus clientes. Pereira (2009) sugere que existem quatro formas para reduzir a assimetria de informação e o risco de crédito, a informação oriunda das agências de referência, os modelos de notação de risco, o preço baseado no risco e as garantias.

A seleção adversa, intimamente ligada com a assimetria de informação, aparece ainda antes de existir uma transação (Mishkin, 2004), evidente sempre que se faz uma pior escolha ou decisão do que a seria esperada se não existisse assimetria de informação. A seleção adversa, segundo Mishkin (2004) aumenta a possibilidade de um financiamento ser feito a um mau devedor, ou seja, a um indivíduo com uma elevada probabilidade de incumprimento.

O *creditworthiness* ou capacidade de crédito traduz a disposição e capacidade do devedor honrar os seus compromissos (Anderson, 2007). Segundo Thomas *et al.* (2002) reflete a avaliação do credor ao potencial devedor e traduz as circunstâncias dos mesmos. Esta disposição e capacidade em honrar compromissos são essenciais na avaliação de risco de incumprimento imprescindível para a decisão de conceder crédito.

Por fim, o *moral hazard* surge após a ocorrência de um compromisso (Mishkin, 2004), comporta o risco das partes envolvidas mudarem o seu comportamento de forma indesejável, em função de outros fatores, como leis, situação económica, *etc.*, podendo alterar a probabilidade de incumprimento prevista, numa etapa inicial.

2.1.1 Risco de crédito

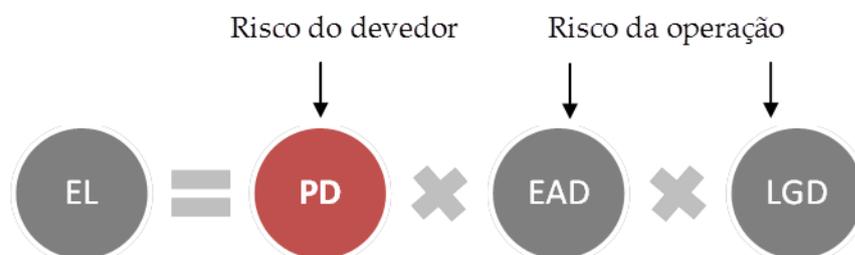
O risco de crédito é o principal risco a que uma instituição de crédito está exposta (Caouette *et al.*, 2008 e Dimitriu *et al.*, 2010) e é consequência das transações financeiras entre investidor e devedor, potenciado pelo crescimento do crédito e sustentado por mudanças económicas, políticas e tecnológicas (Caouette *et al.*, 2008). Conforme enunciado por Paiva (1997), o risco encontra-se implícito em qualquer operação de crédito.

O seu significado prende-se com a possibilidade do devedor não cumprir a sua obrigação, o que terá um efeito adverso para o credor. Segundo Anderson (2007) o risco de crédito comporta o impacto financeiro potencial de qualquer

mudança real ou percebida do devedor. Gitman (2005) afirma que o risco deve ser visto como uma possibilidade de perda financeira, onde os retornos esperados são variáveis. Para Siqueira (2000) o risco de crédito é “consequência da decisão livre e consciente de expor-se a uma situação na qual há a expectativa de ganho sabendo-se que há a possibilidade de perda ou dano”. Maccrimmon e Wehrung (1986) acrescentam ainda que existem três fontes que originam o risco, especificamente a ausência de: controlo, informação e tempo.

Segundo Jorion (2007) e Maccrimmon e Wehrung (1986), o risco de crédito é composto por três fatores de risco; i) o risco de incumprimento, mensurado através da probabilidade de incumprimento (PD); ii) a exposição a risco de crédito (EAD), que mensura o valor em dívida num determinado momento; e iii) o risco de recuperação, que traduz o valor possível de recuperar (LGD), em caso de incumprimento. Com a determinação destes fatores é possível mensurar a perda esperada (EL), como demonstra a Figura 1.

Figura 1 – Cálculo da perda esperada



Fonte: Adaptado de Gaspar (2014)

Enquanto a determinação da EAD e da LGD recai sobre a análise do risco da operação (valor exposto a incumprimento e colaterais), a PD é efetuada por via de uma notação de risco do devedor, principal foco deste estudo. Contudo o risco de crédito não é apenas mensurado pela perda esperada, mas também pela perda inesperada, ou seja, aquela com menor probabilidade de ocorrer e mais difícil de prever (Gaspar, 2014).

Jorion (2007) afirma que é difícil avaliar a probabilidade de incumprimento, dada a sua pouca frequência relativamente ao cumprimento, sendo este o primeiro fator de risco de crédito a que as agências de notação dedicam as suas análises. De acordo com Wesley (1993), são dois os fatores subjacentes ao risco de incumprimento, o primeiro é a pouca qualidade da análise de crédito e o segundo é a deterioração da situação macroeconómica.⁴ O impacto dos eventos sistémicos são sem dúvida variáveis difíceis de estimar e quantificar, já que comportam a ocorrência de fatores externos, não controláveis que podem ou não influenciar as outras variáveis explicativas, daí que não sejam tidos em conta na maioria dos modelos de medição de risco.

Os bancos tentam atenuar o risco de crédito através de várias medidas como o desenvolvimento de notações de crédito assentes em análises rigorosas, a definição de contratos menos flexíveis e um maior enfoque nas garantias reais e financeiras (Anderson, 2007). A existência de sistemas de classificação e análise de risco de crédito, bem geridos promovem a segurança e solidez duma instituição de crédito e facilitam a tomada de decisão (OCC, 2001).

O risco de crédito foi o primeiro a ser alvo de supervisão, por parte dos bancos centrais, através do acordo de Basileia, que estabeleceu requisitos mínimos de capital em função da exposição a risco de crédito. A recente crise financeira despoletou também uma maior atenção ao mercado financeiro, fazendo emergir uma maior regulamentação nesta temática, levando as instituições financeiras a procurem desenvolver os seus próprios modelos de análise de risco de crédito.

⁴ A situação macroeconómica é um fator externo que uma instituição nem sempre consegue controlar e tem dificuldades em prever.

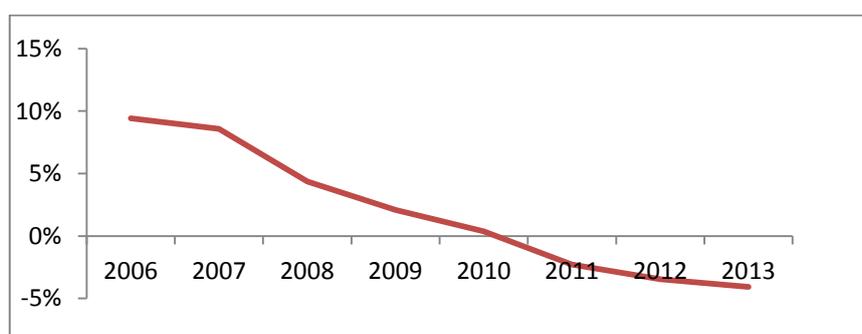
2.1.2 Crédito a particulares em Portugal

O crédito a empresas é um tema amplamente discutido, permanecendo até aos anos 90 relacionado com as crises bancárias, com menor evidência para as perdas na carteira de empréstimos a particulares, que em termos agregados eram pouco significativas. Contudo o aumento do endividamento dos particulares, nos finais dos anos 90, despoletou uma maior atenção sobre esta tipologia (Alves e Ribeiro, 2011).

O maior acesso ao crédito ganhou expressão em Portugal com a desregulamentação do sistema bancário nacional, a liberalização e descida das taxas de juro, e ainda por esta altura, como mencionado por Frade (2007), as instituições bancárias deixaram ter a exclusividade na concessão de crédito a particulares, surgindo outras instituições. Além disso, como referido por Farinha e Lacerda (2010), também a diminuição dos custos de transação propiciaram condições favoráveis ao emergir de novos produtos, que potenciaram a relação entre as famílias e o sistema bancário.

Se até esta data os portugueses apresentavam um baixo nível de endividamento, depressa esse cenário se alterou, fruto da banalização do uso do crédito, indispensável ao quotidiano das famílias, deixando de ser usado em momentos ocasionais de falta de liquidez ou despesas inesperadas, mas passando a ser tido como um meio de pagamento de qualquer tipo de bem ou serviço (Marques e Frade, 2003).

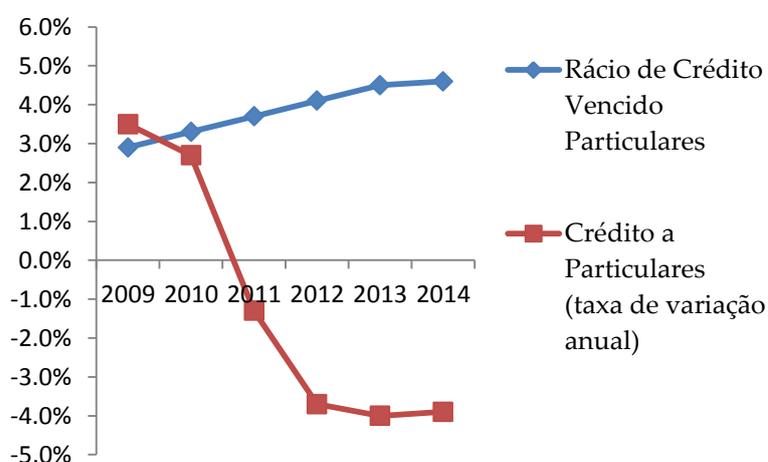
Gráfico 1 – Crédito a particulares em Portugal (variação anual 2006-2013)



Fonte: BdP (2014a)

Segundo dados do Banco de Portugal (BdP, 2014a), o crédito a particulares em Portugal tem vindo a diminuir desde 2007, tendo uma variação anual negativa desde 2011, conforme demonstra o Gráfico 2. Esta contração do crédito, a partir de 2007, é justificada pela crise internacional que teve início nesse mesmo ano. Por outro lado, o ano de 2011 também foi um ponto de viragem na política do país, que se viu obrigado a recorrer a ajuda externa com a dívida pública a atingir os 108,2% do PIB, ao mesmo tempo que se intensificavam as dificuldades em obter financiamento e o desemprego a atingir os 12%, que depressa se refletiu no incumprimento dos compromissos assumidos e no respetivo rácio de crédito vencido que tem vindo a crescer, o que também se veio a espelhar, por outro lado, na diminuição do crédito concedido a particulares.⁵

Gráfico 2 – Rácio de crédito vencido e crédito a particulares (2009-2014)



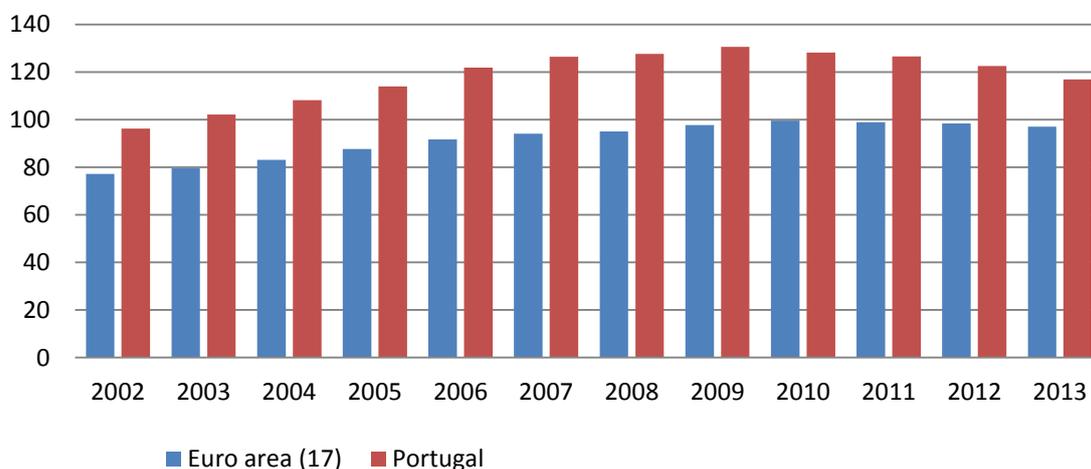
Fonte: BdP: Bpstat, 2014

O rácio da dívida sobre o rendimento das famílias atingiu o seu máximo em 2009, com cerca de 130,5% valor bem acima da média da zona euro, que nesse ano registou 97,6%. A partir de 2009, Portugal tem vindo a apresentar

⁵ Em Maio de 2011 foi assinado o memorando de entendimento onde são enunciadas as condições gerais da política económica, até 2014, sobre a concessão de assistência financeira da União Europeia a Portugal.

sucessivamente um rácio inferior, registando em 2013 apenas 116,8%, embora ainda acima do registado na zona euro, como demonstra o Gráfico 3.

Gráfico 3 – Dívida total em relação ao rendimento das famílias



Fonte: Eurostat, 2014 e BdP (2014a)

O pedido de auxílio externo obrigou Portugal a estabelecer medidas de forma a reduzir a dívida pública, atuando ao nível da melhoria da eficiência da administração pública, da regulação e supervisão do sector financeiro, do mercado de trabalho e educação, do mercado de bens e serviços, do mercado de habitação e do sistema judicial, conforme enunciado no Memorando de Entendimento (2011) entre o Estado Português, o Banco Central Europeu, o Fundo Monetário Internacional e a Comissão Europeia.

Segundo projeções do BdP (2014b) terminado o Memorando a economia portuguesa deverá recuperar de forma gradual a sua atividade, apontando desde logo um crescimento de 1,2% do PIB, para 2014.

2.2 Notação

As raízes da notação remontam a 1936, quando o estatístico Fisher (1936) publicou um artigo sobre análise discriminante linear, para classificar diferentes tipos de íris, servindo de base para muitas outras áreas de estudo (Anderson, 2007), conforme se percebe pela revisão feita por Paliwal e Kumar (2009) desde a saúde, medicina, finanças, contabilidade, engenharia, produção, marketing ou qualquer outra área em que seja importante tomar uma decisão. Durand (1941) faz uso da mesma técnica para distinguir bons e maus negócios, mais tarde também Johnson (2004) analisa a qualidade dos empréstimos, com base em várias variáveis desde a idade, sexo, ocupação, ativos, *etc.*

Qualquer técnica que permita, através de um conjunto de variáveis e a relação entre as mesmas no passado, prever o risco futuro de um devedor é, segundo Schreiner (2000), designado por notação. Podem-se distinguir duas abordagens de notação: a subjetiva, ou a estatística (Schreiner, 2003).

Na notação subjetiva, a análise de risco de crédito parte do conhecimento e experiência do analista e portanto as variáveis consideradas na análise são maioritariamente qualitativas, o que poderá levar a problemas de consistência e coerência, ao não permitir estabelecer um padrão de análise, sendo também mais dispendiosa (Lehmann, 2003).

Por sua vez, a notação estatística tem subjacente a análise estatística de variáveis quantitativas, sendo possível desenvolver e aplicar um modelo *standard* e coerente, que para além de tornar a decisão mais célere, permite que seja também mais assertiva. Daí se compreende que o estudo de Thomas *et al.* (2002) tenha constatado que, aplicando um modelo de notação estatístico, o número de maus pagadores desça mais de 50%.

Qualquer que seja a abordagem adotada, segundo Schreiner (2003), a relação entre variáveis e o risco de crédito será igual no futuro como no passado, o que nem sempre se verifica. É exatamente nesta questão, de usar dados históricos

para prever o futuro, que facilmente se compreende que uma abordagem subjetiva é mais flexível que a estatística, que se mostra mais rígida na mudança. Em países desenvolvidos a abordagem estatística é a mais usada (Mester, 1997), já que reduz o tempo e o custo da análise de risco e uniformiza a respetiva análise, e portanto será esta a abordagem seguida neste estudo.

2.2.1 Notação de risco

No que respeita à notação de risco, segundo Thomas *et al.* (2002), a história do crédito tem cerca de 5.000 anos, enquanto a notação de risco de crédito tem pouco mais que 60 anos, resultado da vulgarização do uso de cartões de crédito e do avanço tecnológico (Anderson, 2007), já que a cada dia era mais inviável tomar decisões de empréstimos que não fossem automáticas, dada a sua procura (Thomas, 2000).⁶

Segundo Hovakimian *et al.* (2009) as notações refletem avaliações "relativas" de risco de crédito, num dado momento do tempo, estando vulneráveis às mudanças externas que possam ocorrer. Espera-se que a notação traduza uma previsão de risco futuro, conforme enunciado por Dimitriu *et al.* (2010). Júlio (2013) afirma que a notação de risco de crédito é uma análise estatística à qualidade do crédito, que correlaciona o incumprimento com as características do devedor. Depressa esta análise foi usada para outro tipo de produtos, como os empréstimos pessoais e à habitação e a pequenas empresas (Anderson, 2007; Saunders e Cornett, 2008; Dimitriu *et al.*, 2010). Paralelamente, cada vez mais as instituições de crédito sentiam uma maior pressão relativamente à concorrência (Hand e Henley, 1997), exigindo-se uma melhoria das receitas e uma procura de

⁶ Cabe ressaltar, que apesar do termo risco ser muitas vezes ligado ao termo incerteza, estes não são sinónimos, como conclui Siqueira (2000) e Paiva (1997), numa situação de incerteza não se conhece a distribuição de probabilidade, o que não acontece quando se menciona o termo risco, em que é possível estabelecer resultados e as respetivas probabilidades associadas.

mecanismos mais eficientes de atrair clientes, diminuir tempo e perdas (Semedo, 2009).

A Fair Isaac foi a originadora do conceito dirigido para particulares. Fundada em 1956, em San Francisco CA, por Bill Fair e Earl Isaac (FICO, 2014b), define notação de crédito como um número que sintetiza o risco de crédito baseado numa imagem, num dado momento, do relatório de crédito de determinado sujeito (FICO, 2005). Muitas foram as críticas, principalmente por parte das seguradoras de crédito que viam o crédito como uma arte, não como uma ciência, séticas relativamente à capacidade de um computador conseguir tomar decisões de crédito (Anderson, 2007).

Tradicionalmente, a determinação da notação de risco dos devedores era feita de acordo com os 5C's do crédito: caráter, capacidade, capital, colateral e condições. Esta avaliação permitiria uma análise histórica e prospetiva do risco de crédito (Santos, 2003 e Anderson, 2007). O caráter traduz a vontade do devedor em cumprir os seus compromissos (Sousa e Chaia, 2000), a capacidade reflete a aptidão para pagar os compromissos assumidos, o capital compreende todos os recursos disponíveis para fazer face a eventos não previstos, o colateral é composto por todos os ativos que podem fazer face ao cumprimento da dívida, e por fim as condições prendem-se com o impacto do ambiente atual no cumprimento do devedor (Anderson, 2007).

Este tipo de análise, de cariz mais subjetivo como é exemplo os 5C's do crédito, só é possível quando se conhece o devedor, mas é ineficiente atualmente, dada a grande mobilidade de pessoas e bens (Anderson, 2007). Não deixa, contudo, de se ter em conta sempre que possível, já que enriquece a análise de risco de crédito.

A FICO (2005) acrescenta que a análise tradicional de risco de crédito torna o processo mais lento, inconsistente e com maior probabilidade de errar. Contudo, este tipo de análise é a génese da banca relacional, que se baseia na confiança e na relação com o devedor. Assim, se inicialmente os modelos

apenas discriminavam entre aceitar ou rejeitar o crédito (Anderson, 2007), mais tarde surgiram os modelos quantitativos ou estatísticos, o que deu forma à banca transacional apoiada nas características dos clientes, utilizando-as para calcular uma notação que represente a probabilidade de incumprimento.

A notação de risco permite estabelecer numericamente quais os fatores de incumprimento, avaliar a importância relativa dos fatores de incumprimento, aperfeiçoar os modelos *pricing* do risco de incumprimento, melhorar a capacidade de monitorizar os clientes de elevado risco e um cálculo mais assertivo de reservas (Saunders e Cornett, 2008).

Atualmente, a notação assume grande relevância e serve como meio de controlo do crédito, sendo determinante para várias áreas da atividade financeira, desde avaliação de risco, receita, retenção, marketing, processamento de novos negócios, gestão de contas, recuperação de crédito, *etc.* (Anderson, 2007).

Apesar da evolução da análise de risco de crédito, esta continua sustentada em informação histórica, contendo informação socioeconómica e relativa à relação com a instituição de crédito. Porém, como salienta Schreiner (2003), a relação entre características do devedor e o risco de crédito nem sempre se repete no futuro, o que torna a análise mais suscetível a erros, sempre as condições não são idênticas ao passado, ao cingir-se a dados históricos deixa-se de considerar a dinâmica própria da atividade financeira, que é fruto da absorção de situações repentinas e muitas vezes imprevisíveis.

Na prática, notação de risco reflete a junção de dois conceitos, crédito e notação, consistindo num processo que transforma as variáveis explicativas consideradas numa notação de risco de crédito, que permitirá servir de apoio à tomada de decisão, por exemplo na concessão de crédito (Anderson, 2007).

Segundo Anderson (2007), as instituições baseiam as suas análises em várias subnotações consoante a fonte de informação, a função ou o que se pretende medir, dando origem ao: i) *application score*, usado para novos negócios que

combina dados do cliente, de relações passadas e de agências de referência; ii) *behavioural score*, usado na gestão de conta (limites, autorizações, etc.); iii) *collections score* usado no processo de cobrança, como forma de guiar o comportamento incorporando dados das agências de referência; iv) *customer score* que combina o comportamento de várias contas, usado na gestão de contas e vendas entre clientes; e por fim v) *bureau score* consiste na notação fornecida por entidades externas tidas como referência, relativamente à determinação da probabilidade de incumprimento.

Das subnotações enunciadas é pertinente aprofundar a diferença entre *application score*, comumente designado apenas notação de risco, e o *behavioral scoring*. Segundo Seneviratna (2006) e Dimitriu *et al.* (2010), a notação de risco aplica-se a um novo candidato a crédito, enquanto o *behavioral scoring* é o conjunto de técnicas que auxiliam decisões para um candidato já existente (tendo por base dados históricos e atuais), para os quais existem mais variáveis à disposição (Thomas *et al.* 2002). Seneviratna (2006) destaca ainda que o *behavioral scoring* é usado para decisões de risco, gestão de conta, cobrança e recuperação, tal como uma forma de seleção das campanhas de marketing, sendo que ambas as notações têm como objetivo prever risco e não o explicar (Thomas *et al.*, 2002).

2.2.1.1. Notação de risco de particulares

Tradicionalmente as notações são dirigidas a pessoas coletivas (estados ou empresas) e a produtos financeiros, contudo a notação de risco a particulares começa a ganhar destaque, sustentada pelo exemplo vindo dos EUA e pela crise financeira atual, que em muito se deveu no erro na análise de risco de particulares. Conforme enunciado por Lyons *et al.* (2007), quando se trata de um *score* atribuído a uma pessoa particular este pode afetar as decisões dos credores na concessão de crédito, na capacidade da pessoa para conseguir um emprego,

na compra ou arrendamento de uma casa ou na compra de um carro. Assim quanto melhor for o *score*, menores serão as taxas de juro e garantias exigidas. Se um devedor é considerado de alto risco de crédito, menor o seu *score*, sendo várias as medidas que podem ser tomadas, de forma a diminuir potenciais perdas (EL), como a diminuição do crédito e/ou serviços disponíveis, a atribuição de um limite de crédito, a requisição de determinado valor em depósitos, a cobrança de taxa de juro superior e/ou maior prémio em apólices de seguro (Siddiqi, 2006). Todas estas medidas devem ser acompanhadas por um constante controlo do devedor, para facilmente e rapidamente serem identificados comportamentos indesejáveis.

Sempre que se determinam notações para particulares é essencial que as pessoas as saibam interpretar. Lyons *et al.* (2007) avaliaram o conhecimento nesta matéria, nos EUA, através de 23 questões, concluindo que a maioria dos consumidores compreende o essencial das informações de crédito, mas apresentam algumas falhas sobre os comportamentos que podem afetar a notação, o estudo demonstrou que é premente a aposta na formação, para que se conheçam as vantagens e desvantagens da atribuição de uma notação de risco.

A notação de risco de particulares ganha forma através do *scorecard* ou cartão de notação de risco, como acontece nos EUA. O que comum se designa de *scorecard*, nada mais é do que um bilhete de identidade de risco de crédito de um determinado devedor que tem por base um conjunto de variáveis consideradas explicativas. Segundo Siddiqi (2006), estas variáveis podem ser de várias ordens, desde dados demográficos, relativas à relação existente com a instituição, informação de agências de referência, dados do setor imobiliário, *etc.* Desta forma, a cada variável considerada é atribuída determinada pontuação (*score*) com base em análises estatísticas que têm em consideração resistência, correlação e outros fatores, assim o *scorecard* é a soma dos *scores* atribuídos a cada variável.

2.2.1.2 Vantagens e obstáculos às notações de risco

A notação de risco traz consigo vantagens, mas também alguns obstáculos ao seu desenvolvimento. A notação de risco não garante que não ocorram erros de classificação dos devedores (Investe, 2014). Isto facilmente se percebe, dado que nenhum sistema de análise de risco de crédito consegue todas as variáveis relevantes que caracterizam um devedor (Sousa e Chaia, 2000), seja por implicar demasiados custos ou pela proteção dada ao próprio devedor. A análise e respetivo resultado, que sustenta a notação de risco, é o reflexo da multiplicidade e qualidade da informação e do modelo usado.

As vantagens da notação de risco, segundo diversos autores (Mester, 1997; Brill, 1998; Avery *et al.*, 2000; West, 2000; OCC, 2001; Schreiner 2000, 2004 e 2005 e Dimitriu *et al.*, 2010), incluem:

- a) a redução do custo de análise de crédito, permitindo decisões de crédito mais céleres e transparentes;
- b) um acompanhamento mais próximo;
- c) uma melhor definição de prioridades;
- d) a redução da sinistralidade, ou seja o incumprimento do devedor;
- e) maiores exigências para notações que tenham subjacentes maiores probabilidades de incumprimento;
- f) a aplicação universal e uniforme;
- g) maior facilidade de compreensão e aplicação;
- h) melhor conhecimento da carteira de clientes;
- i) melhor identificação e controlo das variáveis de risco, tal como a previsão de provisões a reconhecer;
- j) a redução da assimetria de informação;
- k) melhor estipulação do *pricing* (em função do risco);
- l) a redução do problema de seleção adversa e
- m) um modelo capaz de se adaptar a nova informação aumentando a sua precisão.

Os obstáculos que podem surgir ao desenvolvimento de notações de risco, segundo alguns autores (Mester, 1997 e Schreiner 2000, 2004 e 2005), incluem:

- a) uma maior exigência em termos de quantidade e qualidade de informação para que o modelo seja eficaz;
- b) a sensibilidade do modelo à realidade da informação que o compõe;
- c) a necessidade de ser assimilado pelos colaboradores da instituição, pois depende da participação de todos, não devendo substituir plenamente outros meios de gestão de risco, como é exemplo a análise fruto da experiência do analista;
- d) como medida de probabilidade não é imune a erros, e
- e) algumas críticas no que toca às variáveis a incluir no modelo (em termos morais e/ou legais).

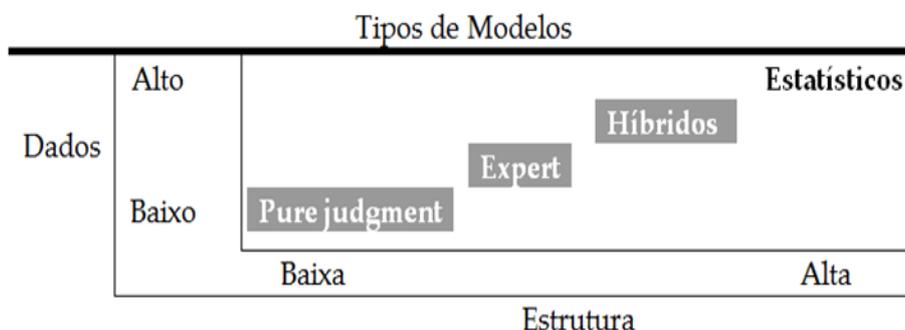
Depois de exposto o seu significado, as suas vantagens e obstáculos cabe esclarecer quais os modelos estatísticos implícitos à notação de risco tema abordado no tópico seguinte.

2.3 Modelos de notação de risco

Os modelos de notação de risco de crédito surgem da necessidade de quantificar o risco de determinado título ou emitente e têm subjacentes indicadores essenciais para determinar a probabilidade de incumprimento que se traduzirá numa notação (Saunders e Allen, 2010). Os modelos de notação de risco permitem, para além da quantificação de uma notação de risco, reforçar a cultura de crédito, uma melhor definição do *pricing* associado ao risco assumido, melhorar a gestão da carteira de ativos e uma mais eficiente alocação de capital (Caouette *et al.*, 2008).

Os modelos de notação de risco podem ser sinteticamente de 4 tipos consoante a quantidade de dados e a estrutura dos modelos, conforme enunciado por Anderson (2007), e ilustrado na Figura 2.

Figura 2 – Tipos de modelos usados na medição de risco



Fonte: Adaptado de Anderson (2007).

Os modelos *pure judgment* são normalmente designados de modelos subjetivos que se baseiam na experiência dos analistas, não exigindo estrutura ou quantidade de dados. É neste tipo de modelo que se inserem os 5C's do crédito. Os modelos *expert* são também modelos pouco exigentes em relação à quantidade de dados, contudo a experiência dos analistas é suficiente para se definirem regras, que auxiliam a tomada de decisão. Os modelos híbridos são modelos que combinam a experiência dos analistas e análises estatísticas, ou seja é uma combinação dos modelos *expert* e dos modelos estatísticos. Por fim, os modelos estatísticos são os mais exigentes, em termos de estrutura e de quantidade de dados, considerados mais confiáveis que os restantes modelos (Anderson, 2007), e como tal serão o foco deste estudo.

Semedo (2009) e Dimitriu *et al.* (2010) destacam genericamente as principais vantagens dos modelos estatísticos relativamente aos não estatísticos, conforme mostra a tabela seguinte.

Tabela 1 – Vantagens genéricas dos modelos estatísticos de notação de risco

	Menor	Maior
Tempo da análise	•	
Custo	•	
Consistência e precisão		•
Lucro		•
Personalização do atendimento		•
Atualização e disseminação da estratégia de risco e de crédito		•
Qualidade do serviço prestado		•
Compreensão do processo e das variáveis da análise		•

Fonte: Adaptado de Semedo (2009) e Dimitriu *et al.* (2010).

As vantagens explícitas na Tabela 1 sustentam o uso dos modelos estatísticos em detrimento dos não estatísticos.

2.3.1 Modelos estatísticos

As técnicas estatísticas usadas na determinação da notação de risco têm-se multiplicado à medida que também o sistema financeiro se torna mais robusto e exigente, permitindo reduzir custos e tempo de decisão (Mester, 1997).

Caouette *et al.* (2008) afiançam que os modelos de notação de risco não são exclusivamente usados para decisões de concessão de crédito, mas servem também como auxiliar na definição, de montantes de crédito, de *pricing* e estratégia, criando uma linguagem universal de medição de risco e servindo como meio de alerta de potenciais problemas. Uma das grandes preocupações na construção dos modelos de notação de risco de crédito prende-se com a conceção e calibração com base em devedores cumpridores (Banasik e Crook, 2007). Contudo, para um bom modelo, são imprescindíveis as características dos considerados maus devedores.

Os modelos de notação de risco têm por objetivo distinguir os candidatos a crédito consoante o risco, ou seja, consoante a probabilidade de cumprir os seus compromissos. Obrová (2012) salienta que os modelos de notação de risco "espremem" todos os riscos num único número. Esta busca da melhoria dos modelos parte da necessidade de aumentar a precisão das decisões de crédito, com base na notação de risco encontrada, tal como afirma e explora West (2000).

Apesar da sua curta história muitos já foram os modelos usados na determinação de notação de risco para particulares, embora a maioria já tivessem sido testados na notação de estados e empresas. Não há consenso quanto ao melhor modelo a aplicar, já que para todos eles é possível destacar pontos fortes e pontos fracos. A escolha do modelo dependerá das características do problema, ou seja, da estrutura de dados, das variáveis e do objetivo da notação, conforme enunciado por Hand e Henley (1997). Os modelos comumente usados são os de análise discriminante (AD), a regressão logística (RL) e probabilística (RP), as árvores de decisão (RT) e as redes neuronais (RN) (Anderson, 2007).

West (2000) divide os modelos estatísticos em dois tipos, os paramétricos (AD, RL e RP) e não paramétricos (RN e RT). O que distingue estes modelos é que os primeiros assentam na premissa de que a amostra segue uma distribuição normal, as variâncias são homogéneas, e os dados independentes, sendo que os parâmetros que definem a amostra são a média e o desvio-padrão, por outro lado, os modelos não paramétricos não assentam no tipo de distribuição da amostra, não sendo modelos tão robustos como os paramétricos.

Os modelos enunciados não são os únicos desenvolvidos e estudados em matéria de notação de risco, mas são os mais populares. Como exemplos de modelos alternativos tem-se o *genetic programming* (Ong *et al.*, 2005), o *rough sets* (Ong *et al.*, 2005), *data envelopment analysis* (Min e Lee, 2008), o *nearest-neighbor* (Hand e Henley, 1997), entre outros.

2.3.1.1 Análise Discriminante

O modelo de análise discriminante (AD) foi um dos primeiros modelos desenvolvidos, proposto por Fisher (1936), e como tal é um modelo simples que requer dados independentes, normalmente distribuídos e matriz de covariância com variação homogênea (Johnson e Wichern, 2007). Para construir uma função discriminante eficaz é necessário considerar as relações entre as variáveis e os diferentes grupos (podendo estes ser ou não linear) e que as variáveis não consideradas relevantes possam ser excluídas do modelo, afim de uma maior precisão (Ong *et al.*, 2005).

São várias as críticas, segundo West (2000), dirigidas ao modelo de análise discriminante, desde a natureza categórica, a matriz de covariâncias única (bons e maus devedores), até à consideração de uma distribuição normal (Eisenbeis, 1977 e Jorion, 2007), embora Reichert *et al.* (1983) não considerem esta última como uma verdadeira limitação. Saunders e Cornett (2008) acrescentam que este modelo erra ao considerar os pesos das variáveis constantes, ao assumir que as variáveis são independentes, ignoram variáveis importantes, mas difíceis de quantificar (como por exemplo a reputação do candidato).

2.3.1.2 Regressão Logística e Probabilística

O modelo de regressão logística (RL) é considerado, por West (2000), Yobas *et al.* (2000) e Ong *et al.* (2005), como o modelo mais preciso e adequado para problemas de análise de crédito, de entre os modelos tradicionais, e até mesmo uma boa alternativa aos modelos de redes neuronais (West 2000).

Os modelos RL foram também dos primeiros a ser usados, são modelos simples, de resposta binária e de fácil interpretação, alcançando em 1980 maior popularidade que a análise discriminante (Thomas *et al.*, 2002), por se demonstrar tão eficiente e preciso quando a análise discriminante (Harrell e Lee, 1985).

Os modelos RL são o suporte para a aplicação de políticas de *credit rationing*⁷ que tendem a ser decisões de aceitação ou rejeição, tal como resulta dos modelos binários. Este tipo de políticas são tipicamente usadas no crédito ao consumo, crédito por regra menos complexo (Anderson, 2007).

Os modelos de regressão probabilística (RP) são muito semelhantes aos logísticos, a diferença está no *output* gerado que traduz, não numa resposta binária, mas uma probabilidade.⁸

2.3.1.3 Árvore de Decisão

O modelo designado de árvore de decisão tem implícita uma hierarquia de nós de decisão construídos a partir de informação histórica, importantes para determinar o custo/benefício de integrar mais variáveis no modelo ou mesmo para averiguar sobre a qualidade creditícia do tomador de crédito (Anderson, 2007).

2.3.1.4 Redes Neurais

O modelo de redes neuronais (RN) tem como inspiração o cérebro humano (Lee *et al.*, 2002) no que diz respeito à aquisição de conhecimento e tal como o cérebro humano possui a capacidade de aprender, o que permite que através de constantes interações seja atribuída a notação ideal a cada indivíduo, ou pelo menos uma melhor tradução do seu risco implícito. Segundo Hand e Henley (1997), este modelo é apropriado quando existe uma má compreensão da estrutura de dados. Na prática como o próprio nome indica este modelo funciona como uma rede constituída por várias ligações, os nós que constituem

⁷ *Credit rationing* consiste em restringir a quantidade de empréstimo disponível por indivíduo (Saunders e Cornett, 2008).

⁸ Ou seja, um valor contínuo de 0 a 1

a rede podem ser divididos em três camadas: entrada, oculta e de saída (Lee e Chen, 2005; Ong *et al.*, 2005 e Chuang e Lin, 2009).

Haykin (2001) enumera as particularidades do uso de redes neuronais, desde: i) a não linearidade; ii) o mapeamento de entrada e saída; iii) a adaptabilidade; iv) a resposta a evidências; v) a informação contextual; vi) a tolerância a falhas; vii) a implementação em larga escala e viii) a uniformidade de análise.

Os resultados, da aplicação deste modelo, são comparados com os modelos tradicionais, como a análise discriminante linear, a regressão logística ou as árvores de decisão (West, 2000). Contudo um dos problemas é a interpretação dos resultados e a importância das variáveis consideradas (Lee e Chen, 2005 e Chuang e Lin, 2009). Davis *et al.* (1992) acrescentam que pode demorar mais tempo para treinar, quando comparado com outros modelos, sendo ainda menos flexível em caso de mudança de condições, conforme salientam Saunders e Allen (2010).

2.3.1.5 Comparação entre os vários modelos estatísticos

Finda a caracterização dos principais modelos estatísticos usados no desenvolvimento de notações de risco de incumprimento, é pertinente destacar as vantagens e limitações destes quando comparados.

Tabela 2 – Vantagens comparativas dos vários modelos estatísticos de notação de risco

	AD	RL	RP	RT	RN
Fácil compreensão e aplicação	•	•	•	•	
Autoaprendizagem					•
Resolução de problemas de complexos					•
Interação com ambiente externo					•
Mais flexíveis					•

Como denota a Tabela 2 os modelos de redes neuronais mostram-se melhores ao nível da autoaprendizagem, na resolução de problemas complexos, na interação com o ambiente e mais flexíveis que os restantes modelos, contudo não são de tão fácil compreensão e aplicação.

Os modelos estatísticos apresentam algumas limitações, daí que não seja fácil a escolha do (s) método (s) a usar na determinação de *scores* de crédito.

Tabela 3 – Limitações dos vários modelos estatísticos de notação de risco

	AD	RL	RP	RT	RN
Discriminação de casos extremos		•			
Coefficientes estimados das variáveis explicativas mantêm-se constantes	•	•	•	•	
Dados históricos para prever o futuro	•	•	•	•	
Inexistência de razões económicas na escolha de variáveis explicativas	•	•	•	•	•
Escassez e qualidade dos dados	•	•	•	•	•
Custo e demora das bases de dados	•	•	•	•	•
Alteração da cultura da organização	•	•	•	•	•
Difícil compreensão de como a relação entre as variáveis explicativas é estimada					•

Fonte: Adaptado de Semedo (2009).

Das limitações elencadas na Tabela 3, quatro delas são transversais a todos modelos estatísticos considerados, como a inexistência de razões económicas na escolha de variáveis explicativas, a escassez e qualidade dos dados, o custo e demora das bases de dados e alteração da cultura da organização, estas limitações devem ser minimizadas, através da calibragem e validação dos modelos, de forma a não ser posta em causa os resultados do modelo utilizado.

2.3.2 Etapas dos modelos de notação de risco

Diversas são as etapas necessárias para desenvolver um modelo de notação de risco, desde a seleção das variáveis, a determinação de variáveis significantes, a estimação dos coeficientes para cada variável, a estimação de probabilidade de incumprimento, a validação e calibragem do modelo.

Gonçalves *et al.* (2013) enumeram sete passos necessários para a construção de um modelo de notação:

- i) recolha de uma base histórica de clientes;
- ii) agregação dos clientes conforme o seu padrão de comportamento e definição da variável resposta (binária/probabilística);
- iii) seleção de amostra representativa;
- iv) análise das variáveis segundo os critérios estipulados;
- v) escolha e aplicação do tipo de modelo de notação;
- vi) definição de critérios de validação; e
- vii) seleção e implementação do modelo adequado.

Saunders e Cornett (2008) salientam que qualquer que seja o modelo de notação usado deve ser capaz de estabelecer as variáveis relevantes que influenciam o risco de crédito e qual o seu peso, auxiliar na determinação do preço a cobrar para cada nível de risco de crédito e no cálculo de reservas para perdas esperadas, capaz de distinguir bons e maus candidatos.

Para uma instituição financeira recente no mercado o desenvolvimento de um modelo de notação de risco de crédito pode ser inviável, já que para além de ter pouca informação dos clientes, não possui os dados históricos necessários para desenvolver um modelo estatístico (Espin e Rodríguez, 2013). Na maioria das vezes recorrem, portanto a agências de referência para procederem à avaliação do devedor ou como forma de validação dos seus modelos internos de notação de risco. Nem sempre as agências de referência possuem informação sobre determinado candidato, e aí torna-se mais difícil a determinação de uma

notação de risco, Espin e Rodríguez (2013) tentam ultrapassar esta questão ao apresentarem uma metodologia geral para a construção de um modelo de notação, que tem em conta apenas variáveis sociodemográficas.

2.3.2.1 Variáveis do modelo de notação de risco

Uma das etapas iniciais, qualquer modelo escolhido, é a escolha das variáveis. Vários foram os autores que nesta matéria definiram variáveis para estimarem os seus modelos de notação de risco de crédito para particulares.

Comparando os vários estudos, Hand e Henley (1997), West (2000), Lee e Chen (2005), Carvalho (2009), Dimitriu *et al.* (2010), Medina e Selva (2012), Espin e Rodríguez (2013) e Júlio (2013), que usam bases de dados de pessoas singulares (anexo 1) e considerando-se apenas as variáveis que se são mencionadas por mais que um autor concluiu-se que as variáveis mais frequentes são a idade e o estado civil (88%), seguindo-se o prazo do empréstimo e a antiguidade da conta bancária (63%) e as variáveis anos na residência atual e anterior, valor do empréstimo, anos empregado na empresa atual, rendimento, tipo de propriedade (habitação própria/arrendada) e género mencionadas por 50% dos autores mencionados.

Por outro lado, outros autores usam à semelhança deste estudo bases de dados relativas a famílias, como é o caso de Canner *et al.* (1991), Godwin (1999), Zhang e DeVaney (1999), Stavins (2000), Getter (2003), Duygan e Grant (2009), Alfaro e Gallardo (2012) e Costa (2012a). Apesar de usarem o mesmo tipo de dados os autores vão divergindo também relativamente ao conjunto de variáveis que utilizam para determinar a probabilidade de incumprimento (anexo 2). Uma vez mais, considerando apenas as variáveis que são mencionadas por mais que um autor, concluiu-se que a variável idade foi incluída por todos os autores, que a variável estado civil é incluída em 88% dos estudos, seguindo-se a inclusão da variável relativa à dimensão da família ou número de crianças do agregado (75%), com uma frequência de 63% surgem as

variáveis rendimento e situação laboral, e com 50% as variáveis propriedade de habitação, riqueza, escolaridade e raça.

Em suma, esta análise comparativa dos estudos realizados em matéria de notação de risco, permite concluir que algumas variáveis são comumente incluídas, como são exemplo a idade, o estado civil, o rendimento e a propriedade de habitação.

Nem sempre se consegue todas as variáveis que seriam ideais (Medina e Selva, 2012), mesmo que segundo os autores Abdou e Pointon (2011) e Greene (1992) não exista um número ideal de variáveis para a análise de risco, seja devido ao modelo de notação de risco usado, seja por questões de proteção de informação do particular que por regra se traduzem na legislação em vigor (Hand e Henley, 1997 e Thomas *et al.*, 2002) que podem restringir a inclusão de certas variáveis, como raça, religião, género, *etc.* ou até mesmo culturalmente inaceitáveis, como salienta Thomas *et al.* (2002), como é exemplo o historial de saúde, sendo imprescindível que as variáveis estejam correlacionadas com o risco de incumprimento (Dimitriu *et al.*, 2010).⁹

A falta de informação desejada para avaliar a qualidade do devedor, pode provocar um enviesamento da avaliação (mais penalizadora), contudo pode levar o devedor a ter um maior incentivo para fornecer todas as informações necessárias (Lehmann, 2003). A escolha da informação necessária é, segundo Sicsú (2010), o segredo para a obtenção de um bom modelo.

2.3.3 Breve nota aos estudos realizados em Portugal

No que diz respeito a uma notação dirigida a particulares, em Portugal, esta ideia já foi abordada por alguns autores onde se destacam Silva (2011) e Pereira (2011) inspirados no modelo americano (FICO). Silva (2011) apresentou um

⁹ Por exemplo no Reino Unido não é permitida a inclusão da variável género e raça, conforme a legislação em vigor (Hand e Henley, 1997).

modelo de *scoring* para microcrédito, por sua vez Pereira (2011) vai mais longe ao aplicar um inquérito de opinião a particulares sobre a sua vontade em aderir a um serviço que analisasse o historial de crédito de forma rigorosa e facilitasse a sua relação com a banca, o qual apresentou um elevado número de respostas afirmativas (71%), desta forma traduz a viabilidade da aplicação de produtos de pontuação de risco de particulares em Portugal e daí a pertinência do presente estudo.

Um problema transversal a qualquer modelo prendesse com os dados que o compõe, conforme afirmação de Augusto Castelo Branco¹⁰, “em Portugal, não existe menos informação do que noutros países. Está é menos organizada e com deficiente tratamento informático” (Cordeiro, 2007). Albano Santos¹¹ acrescenta que os bancos cometem um erro grave ao não disponibilizarem os dados a quem os sabe tratar (Cordeiro, 2007). Desta forma a existência de uma notação dirigida a particulares possibilitaria, conforme já mencionado, melhorar e acelerar o processo de concessão de crédito, nesta que é uma altura de profundo interesse e preocupação na avaliação de risco de crédito.

Outros autores têm publicado no Relatório de Estabilidade Financeira (REF), publicado pelo BdP, vários artigos ligados à medição da probabilidade de incumprimento de particulares em Portugal. Farinha e Lacerda (2010) debruçaram-se sobre a relação entre os diversos tipos de crédito e a situação de incumprimento utilizando saldos de crédito, de todas as famílias, inscritos no sistema bancário português¹², estimando um modelo de regressão logística para cada um dos segmentos de crédito.¹³ Concluindo que os detentores de crédito à habitação têm menor probabilidade de entrar em incumprimento que os

¹⁰ Diretor da Informa DB

¹¹ Diretor da Credinformações, única empresa privada que agrega dados de particulares em Portugal

¹² Dados contidos na base da CRC.

¹³ Os segmentos considerados: crédito à habitação, crédito automóvel, dívidas associadas a cartão de crédito, crédito ao consumo e outros créditos.

restantes segmentos, também sugerindo, por outro lado, que a probabilidade de incumprir é maior: i) no crédito conjunto do que no individual; ii) quanto maior o número de bancos credores num mesmo segmento; iii) quanto maior o número de empréstimos; iv) quanto maiores os encargos com dívida; v) quanto maior a maturidade do empréstimo no crédito à habitação; vi) quanto maiores as garantias pessoais; vii) quanto menor for a idade do devedor e viii) se o devedor exercer atividade em nome individual.

Mais tarde também Alves e Ribeiro (2011) desenvolveram dois modelos econométricos¹⁴, usando dados agregados relativos a fluxos de empréstimos vencidos e de cobrança duvidosa e variáveis explicativas macroeconómicas, como a taxa de juro a 6 meses, *spread*, variação homóloga do PIB e do desemprego e uma *dummy* para o período da crise, isto no que diz respeito ao segmento de empréstimos à habitação. No segmento dos empréstimos ao consumo e outros fins usou as mesmas variáveis à exceção da *dummy*. Os modelos desenvolvidos pelos referidos autores mostraram-se suficientemente robustos para prever a evolução dos empréstimos vencidos e de cobrança duvidosa em diferentes cenários macroeconómicos.

À semelhança do presente estudo outros autores usaram nas suas análises dados do Inquérito à Situação Financeira das Famílias (ISFF) realizado em 2010. Costa e Farinha (2012c) desenvolveram uma análise microeconómica com base nos resultados que investiga o efeito de um conjunto de características económicas e sociodemográficas das famílias sobre a probabilidade das famílias participarem no mercado de dívida, concluindo que a participação no mercado de dívida aumenta: i) com o nível de escolaridade; ii) com o rendimento; iii) quando o indivíduo de referência é trabalhador por conta de outrem e iv) com o número de crianças ou número de elementos do agregado.

¹⁴ Um para cada segmento considerado relevante dos empréstimos a particulares: habitação e consumo e outros fins.

Costa (2012a) utilizou também o ISFF de 2010 para estimar uma probabilidade de incumprimento das famílias com base na ocorrência de atraso ou incumprimento nos últimos 12 meses, na influência da alteração das condições das famílias e no tipo de empréstimo. Concluindo que cerca de 12% das famílias com empréstimos falhou ou atrasou o pagamento nos últimos 12 meses, tendo a proporção de incumprimento sido menos expressiva para maiores níveis de riqueza e rendimento. E que as famílias em que tenham ocorrido alguma alteração à sua situação financeira desfavorável¹⁵ têm maior incidência de incumprimento.

¹⁵ Perda de emprego nos últimos 3 anos, redução da riqueza líquida, menor rendimento que no ano anterior ou aumento das despesas regulares face ao ano anterior.

2.4 Agências de notação de risco

Este subcapítulo tem como intuito distinguir os dois tipos de notação, as notações solicitadas e as não solicitadas, um breve enquadramento aos diferentes tipos de agências de notação de risco, as de *rating*, as de *scoring* e as agências de referência ou *credit bureaus*, e por fim destacam-se as agências de notação em Portugal.

2.4.1 Tipos de notação das agências de notação de risco

As notações desenvolvidas pelas agências podem ser solicitados ou não solicitados, as solicitadas são as usuais atualmente, sendo pagas e solicitadas pelo devedor (Butler e Cornaggia, 2012). Contudo nem sempre foi assim, inicialmente as notações eram procuradas e pagas pelos investidores como forma de auxiliar as suas análises (White, 2010).¹⁶

O facto da maioria das notações serem solicitadas por quem necessita de financiamento é, atualmente, uma das críticas apontadas às notações desenvolvidas pelas agências (Butler e Cornaggia, 2012), ainda mais quando, como menciona Kisgen (2006), as políticas de financiamento assentam na notação que se quer conseguir ou manter, frequente quando se tratam de empresas.

As notações não solicitadas não escapam às críticas, segundo Baker e Mansi (2002), podem levar a que a empresa ou emitente apenas contrate as agências com avaliações favoráveis. Porém não são considerados mais precisas do que as solicitadas (Butler e Cornaggia, 2012), já que as agências não têm acesso a toda a informação, baseando as suas análises em informações públicas. Daí que, Ellis (1998) conclua que é importante, para os emitentes, uma clara identificação das classificações que não são solicitadas. Em sentido contrário, Ellis (1998)

¹⁶ A Egan-Jones Ratings é das poucas agências em que é o investidor que pagar a análise feita pela agência (Saunders e Allen, 2010).

questionou 102 investidores institucionais e conclui que 89% dos mesmos consideram que as notações não solicitadas e solicitadas possuem a mesma qualidade.

As análises realizadas pelas agências de notação servem de apoio na medição do risco, na ótica dos investidores auxiliando-os nas suas análises, sendo assim capazes oferecer serviços de informação de baixo custo que reduzem a diferença entre o que o investidor sabe sobre um devedor (Gavras, 2012).

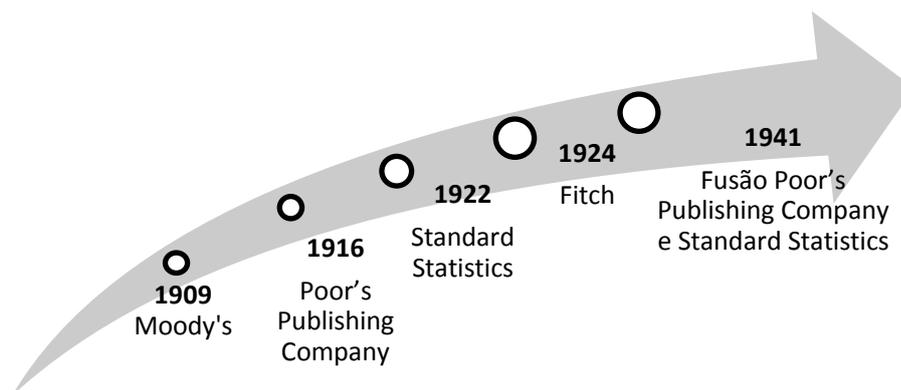
As notações emitidas pelas agências têm vantagens e desvantagens, conforme destaca Smith (2014). As vantagens enunciadas vão, desde um maior incentivo à obtenção de uma melhor notação, na medida em que uma melhor notação propicia menores custos de financiamento, uma melhor relação risco-retorno, possibilitando uma maior validação do nível de risco do emitente e uma fácil compreensão. No que diz respeito às desvantagens, o mesmo autor destaca que nem sempre as notações são precisas, sendo este um ponto valorizado pelos emitentes (Baker e Mansi, 2002), podendo gerar conflito de interesses e ainda a não comparabilidade das notações, justificada pela existência de vários modelos de determinação das notações de risco das várias agências e à variedade de informação, que pode ou não ser incluída na análise.

2.4.2 Tipos de agências de notação

As agências de notação de risco podem assumir 3 denominações distintas, de *rating*, de *scoring* ou de referência, consoante o tipo de notações que emitem e os emitentes a que dedicam as suas notações. A história das agências de notação teve origem nos EUA ligada aos caminhos-de-ferro com a publicação de Henry Varnum Poor em 1860 da primeira compilação de informação financeira e operacional dirigida aos investidores, dando mais tarde lugar à análise dos investimentos nos caminhos-de-ferro.

As agências de *rating* ganharam ênfase no sistema financeiro mundial no século XXI (Graça, 2012), surgindo ainda antes da 1ª Guerra Mundial, sendo a Moody's a sua precursora. Em 1929, a Moody's tinha analisado cerca de 50 governos estaduais e locais dos EUA (Feio e Carneiro, 2012) e a Standard & Poor's 21 estados soberanos (Bhatia, 2002).

Figura 3 – Cronograma das principais agências de *rating*



As agências de *rating* surgiram da necessidade de análise de risco de crédito, podendo se afirmar que são um produto da globalização dos mercados financeiros (Pedro, 2012)¹⁷.

Boot *et al.* (2003) afirmam que “as agências de *rating* podem ser vistas como agências de processamento de informações com poder para acelerar a disseminação de informações nos mercados financeiros.” Em termos jurídicos, as agências de *rating* são pessoas coletivas privadas, que têm como função a emissão de notações de risco, conforme definido no Regulamento (CE) n.º 1060/2009. As agências avaliam a qualidade de crédito, dos títulos de dívida emitidos por empresas, instituições financeiras, produtos estruturados e estados, e dos próprios emitentes (Caouette *et al.*, 2008).

A primeira agência de notação de risco dirigida a particulares surgiu em 1989 com assinatura da *Fair Isaac Company* (FICO), à semelhança das agências de

¹⁷ Prefácio de Guerreiro, P. (2012)

rating, as agências de *scoring* dedicam-se à análise da qualidade de crédito não de pessoas coletivas, mas de pessoas singulares.

Atualmente a FICO é considerada uma das 25 melhores empresas de tecnologia de ponta para serviços financeiros pelo 10º ano consecutivo, conforme o FinTech 100 (IDC Financial Insights, 2013).

Genericamente fala-se de FICO score, contudo existem três diferentes notações FICO desenvolvidas pela FICO, uma para cada agência de reporte de crédito de particulares nos EUA, que assumem denominações distintas, BEACON (Equifax); Experian/FICO Risk Model (Experian) e a FICO Risk Score, Classic (TransUnion). A diferença entre estes *scores* está unicamente nos dados que os compõem, específicos de cada agência enunciada, permanecendo os mesmos métodos e rigor, esperando-se que os *scores*, das três agências, para o mesmo sujeito assumam valores próximos, conforme esclarece a FICO (2005).

A FICO não é contudo a única empresa que se dedica a notação de risco dirigida para particulares. Também a *VantageScore Solutions* desenvolveu o seu próprio modelo de notação, com objetivo de suprir as limitações apontados ao modelo desenvolvido pela FICO, como a exigência de uma linha de crédito aberta há pelo menos seis meses.

A grande diferença entre ambas prende-se com o tratamento dos dados, o histórico de crédito e os tipos de crédito. O modelo VantageScore surgiu em 2006 fruto da vontade da Equifax, Experian e TransUnion em criarem uma solução mais avançada de notação de crédito, com critérios e ponderações diferentes, tal como uma maior padronização comparativamente ao FICO score.

Uma das vantagens apontadas ao VantageScore é a capacidade de fornecer uma notação para um grande número de consumidores, que anteriormente eram excluídos por insuficiente histórico de crédito (Burns, 2010). A última versão do modelo VantageScore 3.0 surgiu da necessidade do mesmo ser adaptado, resultado da crise de 2007, esperando-se que esta última atualização aumente a precisão do modelo anterior, pois possui uma base de dados de

maior qualidade e quantidade. Este modelo inclui ainda outro tipo de dados como de arrendamento, de serviços públicos e de telecomunicações (VantageScore, 2013).

A notação desenvolvida pela FICO divide as variáveis em cinco grupos, a VantageScore 3.0 (VantageScore, 2013) distingue seis grupos, dando-lhes também ponderações distintas, como resume a Tabela 4.

Tabela 4 – Componentes e pesos dados às variáveis do FICO score *versus* VantageScore

FICO score		VantageScore 3.0	
Histórico de pagamentos	35%	Histórico de pagamentos	40%
Montante em dívida	30%	Montante em dívida relativamente ao limite de crédito	20%
Histórico de cumprimento de crédito	15%	Antiguidade, tipos de crédito e respetivos históricos	21%
Outros fatores (tipos de crédito e quantidade)	10%	Saldos atuais e históricos	11%
Créditos recentes	10%	Créditos recentes	5%
		Valor de crédito disponível	3%

Fonte: Adaptado de FICO (2005) e Vantagescore (2013)

As principais diferenças entre as duas notações, FICO e VantageScore 3.0, prendem-se com os pesos das variáveis que as compõem, mas também com as variáveis que incluem. O FICO score considera o montante em dívida com um peso de 30%, enquanto o VantageScore 3.0 pondera o montante de dívida relativamente ao limite de crédito do devedor, com um peso inferior ao FICO score (20%). Também os créditos recentes têm uma maior ponderação na notação FICO (10%) do que na VantageScore 3.0 (5%). Destaca-se, por fim a

inclusão da variável valor de crédito disponível (3%) no VantageScore 3.0 não sendo esta uma variável incluída na notação FICO.

Em termos de escala das notações, a escala desenvolvida pela FICO vai de 300 a 850 conforme maior ou menor risco de crédito (FICO, 2005), enquanto a notação desenvolvida pela *VantageScore Solutions* inicialmente apresentava-se numa escala de 501 a 990 nas primeiras duas versões (1.0 e 2.0) e de 300 a 850 na sua última versão (3.0), usando nesta última versão a mesma escala da notação da FICO. Apesar de ser um modelo recente o VantageScore tem lentamente encontrando o seu lugar no mercado onde a FICO fora pioneira.

Por fim, as agências de notação conhecidas como agências de referência centralizam informação de crédito dos seus aderentes, aos quais posteriormente disponibilizam um relatório detalhado sobre cada devedor (Pereira, 2009), estas têm como objetivo auxiliar o mercado a minimizar assimetrias de informação, que permite promover a liquidez nos mercados, aumentar a atividade financeira e reduzir custos, conforme enunciado por Gavras (2012). Este tipo de agência funciona com base no princípio da reciprocidade, e por isso só as instituições que cedam informação para a base de dados terão acesso à mesma (Pereira, 2009).

Independentemente do tipo de agência de notação e de a quem estas se dirigem, e assim como o crédito, parte da confiança, a atividade das agências é sustentada pela confiança, característica que sustenta o sucesso de qualquer agência (Pianesi, 2011).

2.4.3 Agências de notação em Portugal

A nível mundial existem cerca de 58 agências de *rating* conforme listagem da ESMA (2014b) e a disponível em DefaultRisk.com (2014) (anexo 3).¹⁸ Permitindo concluir que é na Alemanha e nos EUA que se concentram mais agências de

¹⁸ Agências compostas por várias afiliadas ou associadas foram apenas contabilizadas uma vez.

rating, sendo que em média se dedicam a três tipos de notação, com maior foco para notações de empresas (79%), instituições financeiras (55%) e estados (41%).

Atualmente na Europa, como já mencionado existem 23 agências de *rating* registadas e duas certificadas junto da ESMA¹⁹, onde, segundo dados de 2012, 3 agências – Moody’s, Standard & Poor’s e Fitch – possuem cerca de 87% da quota de mercado, sendo todas elas americanas e dedicando-se, conforme dados de 2013 (ESMA, 2013), a notações de risco de empresas, instituições financeiras, seguradoras, estados, produtos financeiros estruturados e obrigações.

Conforme informação divulgada pela ESMA, em 2013, foi possível compreender quais as agências notações que existiam à data, o tipo de notação a que se dedicavam e quais os países de residência das mesmas (anexo 4). Como tal observa-se que a Alemanha é o país com mais agências de notação, 8 agências que representam 36%, seguindo-se a Itália e o Reino Unido, com 6 agências que representam 27%. Relativamente aos tipos de notações emitidas estas distinguem-se entre as que são dirigidas a empresas, a instituições financeiras, a seguradoras, a estados, a produtos financeiros estruturados e a obrigações, sendo que em média cada agência dedica-se a 3 dos tipos de notação enunciados, com maior expressão para notações de empresas (73%), instituições financeiras (50%) e estados (45%). Como se pode então constatar, na Europa nenhuma das agências se dedica a notações de particulares, diferente com o que acontece nos EUA, onde encontramos a FICO e a VantageScore Solutions.

A primeira e única agência de *rating* portuguesa surgiu em 1988, com a denominação de Companhia Portuguesa de Rating, SA (CPR), sendo a suas notações dirigidas a empresas e instituições financeiras. Em janeiro de 2014, esta reforçou a sua posição internacional ao unir forças com agências de outros

¹⁹ Entidade europeia responsável pelo registo e supervisão permanente das agências de notação de risco.

quatro países – África do Sul, Brasil, Índia, e Malásia – alterando a sua denominação para ARC Ratings. Esta nova agência pretende, segundo a SaeR (2014), tornar-se uma alternativa às três grandes agências de *rating* – Moody's, Standard & Poor's e Fitch – apesar da sua quota de mercado registar, em 2012, apenas 0,04%, segundo publicação da ESMA (2013). A ARC Ratings pretende sustentar as suas análises numa abordagem de várias perspetivas e no conhecimento local, apostando no rigor técnico e de governação e em metodologias de *rating* inovadoras. Paralelamente, ambiciona reconquistar a função primordial de antevisão das agências de *rating*.

Portugal até julho de 2011, com descida do *rating* soberano pela Moody's ao qual se seguiu a Fitch e a Standard & Poor's, pouco ouvira falar ou até mesmo se preocupara com a repercussão da análise realizada pelas agências de *rating*, em parte justificada pela sua influência não ser considerada como relevante ou até mesmo algum descrédito no trabalho desenvolvido pelas mesmas. O facto dos títulos de dívida soberana portuguesa terem sido de classificados com *Junk Bonds*, o mesmo será dizer que foi desaconselhado o seu investimento, dada a sua elevada probabilidade de incumprimento, levou a uma maior atenção relativamente ao trabalho desenvolvidos por estas agências. Pela primeira vez, Portugal sentiu o poder destas agências levando a que rapidamente taxas de juros subissem, obrigando a um reforço dos colaterais nas dívidas contraídas externamente. A banca portuguesa, que até à data encontrava-se perante a minimização do problema gerado pela crise de 2007, volta a ver reduzida a sua liquidez junto dos mercados externos. Depressa o país e as empresas portuguesas, tal como todo o povo português compreendeu da pior forma o poder destas agências.

No que diz respeito a agências de referência conforme enunciado por Riestra (2002) o mercado de crédito português dispõe de duas agências de referência: a Central de Responsabilidades de Crédito (CRC), sob a alçada do Banco de Portugal, e a Credinformações (Equifax Portugal), que resultou de uma *joint-*

venture, em 1995, entre a ASFAC (Associação das Sociedades Financeiras para Aquisição a Crédito) e a Equifax (Pereira, 2011). A CRC é uma agência de referência pública, que centraliza informações de crédito de todas as instituições supervisionadas pelo regulador, sendo obrigatória a sua consulta sempre que se analisa a solvabilidade de um candidato a crédito, conforme o decreto-lei nº133/09. Por outro lado, a Credinformações é uma agência de referência privada que concentra informação de pagamento da área das telecomunicações.

3 Modelo de notação de risco

Para alcançar o objetivo deste estudo, criar um modelo de notação de risco para famílias portuguesas, foram delineados os seguintes objetivos específicos:

- 1) Testar a influência das variáveis explicativas selecionadas – riqueza líquida, dívida, rendimento, idade, escolaridade, duração do emprego atual, deterioração das condições (de emprego, rendimento, riqueza líquida ou despesas), número de crianças do agregado, recusa de crédito e propriedade de habitação – na probabilidade de incumprimento das famílias;
- 2) Validação da qualidade do modelo para prever o incumprimento da família;
- 3) Definição de classes de notação capazes de alocar cada família consoante a sua probabilidade de incumprimento;
- 4) Por fim, como a probabilidade de incumprimento é um dos componentes da perda esperada foi ainda determinada a perda esperada, tendo em conta a amostra considerada.

3.1 Caracterização da amostra

Os dados usados no desenvolvimento de uma notação de risco para famílias em Portugal foram extraídos da primeira edição do Inquérito à Situação Financeira das Famílias (ISFF) realizado em 2010, que resultou do contato de 8.000 famílias e culminou numa amostra final de 4.404 famílias. Relativamente a quem respondeu ao inquérito coube ao agregado indicar o indivíduo de referência como representante (Costa e Farinha, 2012b).²⁰

²⁰ “Indivíduo de referência corresponde ao indivíduo indicado pelo agregado como representante ou ao seu companheiro/marido se o indivíduo de referência for do sexo feminino e tiver um companheiro/marido do sexo masculino que pertença à família” (Costa e Farinha, 2012b).

3.1.1 Caracterização e principais resultados do ISFF 2010

O ISFF, inserido no projeto europeu *Household Finance and Consumption Survey* (HFCS), tem como objetivo recolher informação sobre a situação financeira das famílias dos países da área do euro. Este projeto é coordenado pelo BCE, sendo em Portugal coordenado pelo Banco de Portugal e pelo Instituto Nacional de Estatística.

As questões presentes no ISFF incidem sobre várias temáticas, desde o património das famílias, as suas decisões de consumo e poupança, aspetos demográficos, situação profissional e rendimentos. Estas temáticas dividem-se em dois grupos de variáveis que compõem o inquérito, o das *core variables* (HFCN, 2012a), composto por variáveis relativas ao património das famílias, e as *non-core variables* (HFCN, 2012b), relativas a aspetos adicionais do património, atitudes e expectativas das famílias.

Este inquérito, segundo publicação de Coata e Farinha (2012b), permitiu concluir que o rendimento médio das famílias, em 2009, era cerca de 20 mil euros, que o rendimento aumenta até à classe dos 45-54 anos de idade e reduz-se nas classes seguintes e que quanto maior o nível educacional e o grau de responsabilidades da profissão, maior o rendimento. Relativamente à fonte de rendimento a principal fonte é remuneração do trabalho, em que a remuneração do trabalho por conta de outrem representa mais de 50%.

Em termos de riqueza líquida as famílias possuíam em 2010 em média 150 mil euros onde as famílias com indivíduo de referência pertencente à classe dos 55-64 anos e/ou seja trabalhador por conta própria possuem maior riqueza. Contudo e tal como sucede nos rendimentos, a riqueza aumenta quanto maior o nível educacional e o grau de responsabilidade da profissão. Em termos de riqueza real a residência principal representa cerca de 55% em valor, permitindo concluir também que o valor detido pelas famílias aumenta quando maior o rendimento, a riqueza líquida, o nível de escolaridade e o grau de responsabilidade da profissão.

3.2 Seleção das variáveis

A seleção das variáveis de um modelo econométrico de notação de risco de uma família ou indivíduo, como visto no subcapítulo 2.3.2., não é um tema consensual que depende, num primeiro momento, dos dados disponíveis e ainda das próprias escolhas dos autores que se baseiam nas suas experiências e outras referências. O mesmo acontece neste estudo que está restringido às variáveis inscritas no ISFF de 2010, onde os dados relativos a rendimentos, despesas, património e dívidas dizem respeito a todo o agregado familiar. Contudo para as variáveis sociodemográficas e relativas ao emprego consegue-se um detalhe individual.

Para a seleção das variáveis a incluir no modelo foi imprescindível a comparação das variáveis escolhidas por outros autores incluídos no subcapítulo referido, mas também nos estudos que usaram o mesmo tipo de base de dados, com destaque para o de Costa (2012a).

Como variável dependente a variável escolhida foi a variável que reflete a ocorrência de incumprimento ou mora em empréstimos, nos últimos 12 meses. As variáveis explicativas têm diferentes naturezas, desde variáveis de cariz sociodemográfico até às relativas ao consumo e poupança. Foram selecionadas 10 variáveis explicativas, típicas sempre que se usa uma base de dados resultante de inquéritos às famílias, demais variáveis 6 são quantitativas e 4 qualitativas. As variáveis qualitativas foram adaptadas de forma a uma melhor interpretação, configurando variáveis *dummy*. As variáveis independentes encontram-se sintetizadas na Tabela 5.

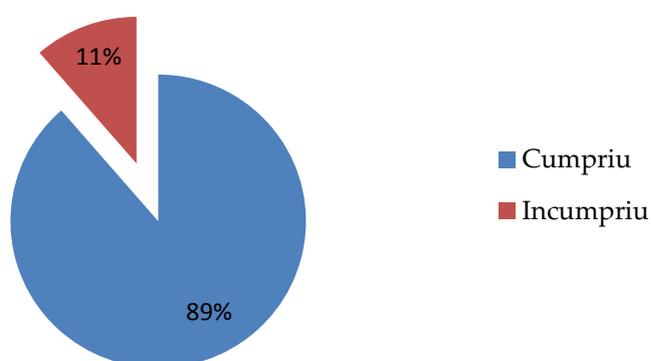
São ainda apresentadas individualmente as principais características das variáveis, consideradas neste estudo, e a respetiva justificação da sua escolha. A Tabela 6, presente no final deste subcapítulo, descreve as principais estatísticas das variáveis selecionadas, características utilizadas para a apresentação individual das variáveis.

3.2.1 Variável dependente

A variável dependente que reflete incumprimento ou mora em empréstimos tem um total de 2.091 observações, num total de 4.404 famílias inquiridas. Como ilustra o Gráfico 4, 89% famílias indicam que cumpriram as suas responsabilidades no ano anterior e apenas 11% indica incumprimento ou falha no pagamento das suas responsabilidades.

Gráfico 4 – Distribuição da variável dependente

Ocorrência de incumprimento ou atraso em empréstimos, nos últimos 12 meses



Fonte: ISFF (2010)

3.2.2 Variáveis explicativas

As variáveis explicativas consideradas são:

- Riqueza líquida;
- Dívida;
- Rendimento;
- Idade;
- Escolaridade;
- Duração do emprego atual;
- Deterioração das condições;
- Número de crianças;
- Recusa de crédito e
- Propriedade de habitação.

Todos os dados estatísticos relativos a estas variáveis são referentes à amostra possível para o modelo observado.

3.2.2.1 Riqueza líquida

O valor de riqueza líquida é relativo ao agregado familiar e resulta da diferença entre o valor de ativos (excluindo pensões públicas e profissionais) e o valor de passivos e outras responsabilidades assumidas pelo agregado. Segundo Costa (2012a) a proporção de famílias com incumprimento apresenta uma tendência de redução acentuada com o nível de riqueza e o mesmo afiança Stavins (2000) ao mencionar que à medida que aumenta o valor de riqueza líquida menor será a probabilidade de incumprimento.

Esta variável está mensurada em euros, tendo sido consideradas respostas de 1.325 famílias, que em termos médios possuem riqueza líquida de cerca de 47 mil euros. Sendo o valor mínimo de riqueza líquida de -363 mil euros, o que indica que a família possui um valor maior de responsabilidades comparativamente ao valor de ativos que detém, por outro lado o valor máximo de riqueza líquida situa-se nos 8 milhões de euros.

Esta variável é uma variável frequente no que diz respeito a amostras como a deste estudo, incluída nos estudos realizados por Zhang e DeVaney (1999), Stavins (2000) e Costa (2012a).

3.2.2.2 Dívida

A variável dívida é referente ao agregado familiar e resulta da junção do valor de dívidas hipotecárias e não hipotecárias. Costa (2012) afirma que é expectável que famílias com níveis de endividamento mais elevados tenham maiores dificuldades em cumprir as suas responsabilidades.

Esta variável está mensurada em euros, obtendo-se resposta para 1.355 famílias, que declaram em termos médios um valor total de dívidas de cerca de

43 mil euros, com um valor máximo de 610 mil euros e mínimo de 10 euros. Esta variável foi também incluída na análise de Costa (2012a).

3.2.2.3 Rendimento

A variável rendimento é referente ao agregado familiar e resulta da junção do valor de rendimentos provenientes do trabalho, de benefícios de desemprego e prestações sociais regulares. Conforme Zhang e DeVaney (1999) e Costa (2012a) a proporção de famílias que incumprem as suas obrigações tende a diminuir com o nível de rendimento.

Esta variável está mensurada em euros, tendo-se obtido resposta para 1.831 famílias, que declaram em termos médios um valor de rendimentos anuais de cerca de 23 mil euros, com um valor máximo de 422 mil euros, onde o valor de rendimentos do trabalho assume maior expressão.

Esta variável foi considerada nos modelos desenvolvidos de Hand e Henley (1997), Lee e Chen (2005), Carvalho (2009), Medina e Selva (2012) e Espian e Rodriguez (2013), mas também por autores com semelhante base de dados como Godwin (1999), Zhang e DeVaney (1999), Stavins (2000), Duygan e Grant (2009), Alfaro e Gallardo (2012) e Costa (2012a).

3.2.2.4 Idade

A variável idade é relativa apenas ao indivíduo de referência do agregado. Costa (2012a) demonstra na sua análise que à medida que a idade aumenta menor o número de famílias com incumprimento.

A idade média do indivíduo de referência da amostra é de cerca de 51 anos, com um valor mínimo de 20 anos e um valor máximo de 85 anos.

Esta é uma variável de aceitação generalizada no desenvolvimento deste tipo de modelo, como demonstrado no subcapítulo 2.3.2 e nos anexos 1 e 2, usada nos modelos propostos por Hand e Henley (1997), Lee e Chen (2005), Carvalho

(2009), Dimitriu *et al.* (2010), Medina e Silva (2012) Espian e Rodriguez (2013) e Júlio (2013), tal como, por Zhang e DeVaney (1999), Canner *et al.* (1991), Godwin (1999), Stavins (2000), Getter (2003), Duygan e Grant (2009), Alfaro e Gallardo (2012) e Costa (2012a), estudos com o mesmo tipo de amostra.

3.2.2.5 Escolaridade

A variável escolaridade é também referente ao indivíduo de referência do agregado. Esta variável foi recodificada de forma a distinguir apenas duas situações: nível de escolaridade superior e restantes níveis de escolaridade. Costa (2012a) conclui que o incumprimento das famílias está associado a menores níveis de escolaridade e que a detenção de um nível de escolaridade superior poderá refletir uma maior capacidade na tomada de decisões de endividamento adequadas.

Conclui-se que cerca de 11% possuem um nível de escolaridade igual ou acima do ensino superior.

Esta é também uma variável comumente considerada como denotam os modelos desenvolvidos por Lee e Chen (2005), Espian e Rodriguez (2013) e Júlio (2013), tal como Costa (2012a) e Getter (2003).

3.2.2.6 Duração do emprego atual

A duração do emprego atual diz respeito ao indivíduo de referência do agregado, sendo expectável que à medida o indivíduo indica mais anos de trabalho na empresa atual, maior a sua estabilidade profissional e financeira como corrobora a análise de Zhang e DeVaney (1999).

Mensurada em anos, foi respondida por 1.330 famílias, o que facilmente se compreende já que estando o indivíduo não empregado esta questão não seria colocada. O valor médio de anos no emprego atual é de cerca de 15 anos, com um valor mínimo de zero e máximo de 45 anos.

Esta variável é também amplamente incluída neste tipo de modelo como denotam os modelos propostos por Hand e Henley (1997), Lee e Chen (2005), Carvalho (2009), Dimitriu *et al.* (2010) e Espian e Rodriguez (2013), tal como o estudo de Zhang e DeVaney (1999) com o mesmo tipo de dados agregados da família.

3.2.2.7 Deterioração das condições

A variável deterioração das condições refere-se ao agregado familiar e reflete a deterioração das condições de emprego, rendimento, riqueza líquida ou despesas. Segundo Duygan e Grant (2009) a ocorrência de mora no incumprimento de obrigações está interligada com alterações desfavoráveis do rendimento ou saúde, também Costa (2012a) conclui que as famílias com alterações desfavoráveis da sua situação financeira apresentam maiores incidências de incumprimento do que as famílias em que essas alterações não ocorreram, justificando a inclusão no modelo proposto.

Esta é uma variável de cariz qualitativo, onde cerca de 63% das famílias declaram ter acontecido, nos últimos 3 anos, algum evento que deteriorou as suas condições.

Esta variável também foi usada por Costa (2012a) e de forma desagregada por Getter (2003) e Duygan e Grant (2009).

3.2.2.8 Número de crianças

O número de crianças refere-se ao número de crianças que compõem o agregado familiar. Segundo Costa (2012a) a proporção de famílias com incumprimento é mais elevada nas famílias com crianças do que quando é composta apenas por adultos.

Em média existe uma criança por agregado, com um mínimo de zero e um máximo de 8 crianças por agregado.²¹ Esta variável foi considerada por Canner *et al.* (1991) e Duygan e Grant (2009) nos seus estudos.

3.2.2.9 Recusa de crédito

Esta variável, recusa de crédito, tem em conta todos os elementos do agregado familiar e o horizonte temporal dos últimos 12 meses. Godwin (1999) demonstrou que acontecimentos passados de crédito, como é exemplo a recusa de crédito, estão associados a uma maior probabilidade de incumprimento.

Esta é uma variável de cariz qualitativo, que contou com a resposta de 653 famílias onde 22% declaram teres-lhes sido recusado crédito nos últimos 12 meses. Godwin (1999) também incluiu a variável recusa de crédito na sua análise.

3.2.2.10 Propriedade de habitação

A variável propriedade de habitação diz respeito a todos os membros do agregado. Getter (2003) salienta que a propriedade de habitação está ligada com o incumprimento, demonstrando que a propriedade de habitação reduz a probabilidade de incumprimento. Concluindo-se, no presente estudo, que 76% dos agregados possuem propriedade de habitação.

A variável propriedade de habitação foi igualmente incluída nos estudos de Zhang e DeVaney (1999), Stavins (2000), Getter (2003) e Duygan e Grant (2009).

²¹ As crianças foram definidas como elementos do agregado com idade inferior a 25 anos, que não estejam empregados e que não sejam o representante da família nem com este tenham uma relação de ascendentes, cônjuges ou companheiros, conforme definido por Costa e Farinha (2012b).

3.2.3 Resumo das variáveis

Como denota a Tabela 5 foram selecionadas dois tipos de variáveis explicativas, as quantitativas e as qualitativas. Para uma mais fácil interpretação dos resultados, algumas variáveis qualitativas foram transformadas em variáveis binárias (*dummy*), evitando-se que determinadas categorias tenham um número reduzido de observações.

Tabela 5 – Caracterização das variáveis selecionadas

Variável	Descrição	Tipo de variável	Mensuração
Variável dependente	Incumprimento ou atraso no pagamento de empréstimos	Qualitativa	0 – Cumpriu; 1 – Incumpriu
Variáveis independentes	Riqueza líquida	Quantitativa	Valor que resulta da diferença do total de ativos pelo total de responsabilidades do agregado
	Dívidas	Quantitativa	Valor de responsabilidades do agregado
	Rendimento	Quantitativa	Valor de rendimentos do agregado
	Idade	Quantitativa	Em anos
	Escolaridade	Qualitativa	0 – Não superior; 1 – Superior
	Duração do emprego atual	Quantitativa	Em anos
	Deterioração das condições	Qualitativa	0 – Não deteriorou; 1 – Deteriorou
	Número de crianças	Quantitativa	Número de crianças do agregado
	Recusa de crédito	Qualitativa	0 – Não recusado; 1 – Recusado
Propriedade de habitação	Qualitativa	0 – Não; 1 – Sim	

Tabela 6 – Estatísticas descritivas das variáveis selecionadas

	Observações	Média	σ	Mínimo	Máximo	Percentis				
						5	25	50	75	95
Incumprimento ou atraso no pagamento de empréstimos	2.091	0,11	0,32	0	1	0	0	0	0	1
Riqueza líquida	1.325	46.754	319.457	-363.000	8.284.100	-87.500	-27.300	750	33.450	285.750
Dívidas	1.355	42.653	52.932	10	610.000	400	5.000	25.000	63.700	146.900
Rendimento	1.821	23.277	25.469	0	422.244	540	9.440	17.195	28.310	65.904
Idade	2.091	51,23	13,70	20	85	30	41	51	60	76
Escolaridade	2.091	0,14	0,35	0	1	0	0	0	0	1
Duração do emprego atual	1.330	14,80	10,47	0	45	1	6	13	22	35
Deterioração das condições	2.091	0,63	0,48	0	1	0	0	1	1	1
Número de crianças	2.091	0,78	0,95	0	8	0	0	0	1	2
Recusa de crédito	653	0,18	0,39	0	1	0	0	0	0	1
Propriedade de habitação	2.091	0,76	0,42	0	1	0	1	1	1	1

3.3 O Modelo

O modelo de risco de crédito desenvolvido utilizou a técnica estatística da regressão probabilística, à semelhança dos modelos desenvolvidos por Duygan e Grant (2009) e Alfaro e Gallardo (2012).

Como exposto no subcapítulo 2.3, relativo aos modelos comumente usados no desenvolvimento de modelos de notação e risco, constatou-se que o modelo logístico é o mais usado. Contudo tendo em conta o objetivo deste estudo, que é o desenvolvimento de uma notação de risco, implica que se tenha não uma resposta do tipo binária, mas que assuma um valor contínuo, por forma a poder ser facilmente alocado a uma escala de notação de risco.

O modelo de regressão probabilística é um modelo estatístico paramétrico, conforme divisão de West (2000), de fácil compreensão e aplicação, que usa dados históricos para prever o futuro. Este tipo de regressão é indicado quando a variável dependente é categórica e assume apenas dois valores possíveis, que neste caso é cumprimento ou incumprimento. Esta técnica visa maximizar a verossimilhança do evento ocorrer, o que se traduz em melhores estimativas dos coeficientes das variáveis explicativas.

O modelo estimado neste estudo é definido por:

$$\begin{aligned} PD_i = Prob (\text{Incumprimento ou mora}_i \\ = \beta_0 + \beta_1 Riqueza_i + \beta_2 Rendimento_i + \beta_3 Idade_i \\ + \beta_4 Escolaridade_i + \beta_5 Duração\ do\ emprego_i \\ + \beta_6 Deterioração\ das\ condições_i + \beta_7 N^o\ de\ crianças_i \\ + \beta_8 Recusa\ de\ crédito_i + \beta_9 Propriedade\ de\ habitação_i \\ + \varepsilon_i) \end{aligned} \tag{3.1}$$

onde i é a família observada, β_0 é a constante do modelo, β_1 até β_9 são os coeficientes das variáveis explicativas e ε é o erro do modelo que não é

explicado pelas variáveis explicativas incluídas. As variáveis são definidas na Tabela 5.

Tratando-se de um modelo probabilístico o seu output será uma probabilidade referente neste caso à ocorrência de incumprimento da família.

Na análise estatística de implementação do modelo à amostra selecionada foram garantidas as seguintes condições:

- i) Inexistência de igualdade das médias das variáveis explicativas para explicarem a variável dependente (*teste t*);
- ii) Variáveis explicativas não são correlacionadas, através do teste de multicolineariedade, garantindo um coeficiente de correlação inferior a 0,6;
- iii) Significância individual das variáveis explicativas, através do teste de *Wald*, de forma a compreender a significância de cada variável para explicar o incumprimento, garantindo-se a rejeição da hipótese nula e servindo de indicador na definição da ordem de inserção das variáveis no modelo.

O modelo proposto teve por base o método *Forward Stepwise (Likelihood Ratio)*, que consiste na junção do método *Forward* e do método *Stepwise*, este método tem subjacente a inserção das variáveis explicativas conforme o nível de significância, ajustado com base no critério de informação de *akaike (AIC)*. Para definir o modelo foram testados 5 modelos, que divergiram relativamente ao conjunto de variáveis consideradas, estes testes permitiram averiguar qual o modelo mais ajustado para explicar o incumprimento das famílias. Para medir o ajustamento e precisão do modelo proposto recorreu-se ao *AIC*, ao cálculo da razão de verosimilhança e ao teste de *Hosmer e Lemeshow*, não descurando o teste de heterocedasticidade do modelo e a análise do pseudo R^2 . Para validar o modelo recorreu-se à matriz de acerto do modelo e à curva ROC.

3.4 Resultado Empírico

3.4.1 Testes empíricos e ajustamento ao modelo

Para averiguar se as variáveis explicativas selecionadas cumpriam as condições enunciadas no tópico 3.3, procederam-se assim a testes comuns para a persecução de um qualquer método econométrico, como o teste de igualdade das médias e o teste de correlação de onde se concluiu que a variável dívidas é a única variável em que a média é estatisticamente semelhante para os diferentes valores da variável dependente considerada, o que justifica a sua não inclusão no modelo final. Ao mesmo tempo, o teste de correlação das variáveis explicativas não demonstrou a existência de elevada correlação entre estas ($\rho > 0,6$).

Procedeu-se ainda a uma análise de multicolinearidade e de significância das variáveis explicativas. A análise de multicolinearidade demonstra não haver colinearidade. Por seu turno, a análise de significância das variáveis explicativas indica que as variáveis recusa de crédito, deterioração das condições, número de crianças do agregado, duração do emprego atual e riqueza líquida são significativas a 5% justificando a sua inclusão no modelo, por outro lado as variáveis propriedade de habitação, rendimento, idade e escolaridade apesar de não significativas ao nível de 5% foram incluídas no modelo por contribuírem para um melhor ajuste (AIC), como se demonstrará adiante nesta exposição.

3.4.2 Construção do modelo

Segundo Stein (2007), calibrar o modelo traduz a capacidade do modelo prever a probabilidade à priori, para isso neste estudo recorreu-se à análise de estudos similares, por forma a justificar a escolha das variáveis explicativas e para prever e compreender o seu efeito na variável dependente.

A base de dados do ISFF é composta por 5 bases de imputação, que visam suprir respostas do tipo “não sei ou não resposta”. Optou-se por utilizar a base sem qualquer tipo de imputação e sem qualquer tipo de ponderação, já que os resultados não se mostraram significativamente distintos entre as diferentes bases, como se pode constatar no anexo 4.

Para definir o modelo foram testados 5 modelos que divergem quando ao conjunto de variáveis que integram. Seguidamente serão apresentadas as diferenças entre os vários testes, ao nível da inclusão ou exclusão das variáveis explicativas e posteriormente os resultados esperados, com base em estudos similares, e efetivos de cada variável explicativa na determinação da probabilidade de incumprimento das famílias.

As tabelas que se apresentam em seguida, Tabelas 7, 8 e 9, contêm o valor dos coeficientes das variáveis (β) e o respetivo valor do desvio-padrão (valor entre parênteses) e a respetiva significância, onde * representa $p < 0,1$, ** $p < 0,05$ e *** $p < 0,01$.

Tabela 7 – Regressão probabilística para probabilidade de incumprimento das famílias (Modelo 1)

Regressão probabilística para probabilidade de incumprimento	
	Modelo 1
Riqueza líquida	-7,7E-07 (6,5E-07)
Dívida	5,1E-07 (1,0E-06)
Rendimento	-6,8E-06 ** (2,9E-06)
Idade	-0,0123 *** (0,0047)
Escolaridade	-0,1069 (0,1550)
Situação profissional	0,1291 (0,1172)
Deterioração das condições	0,6654 *** (0,1160)
Constante	-0,9141 *** (0,2504)
Número de observações	1.247
Log	-443,89
AIC	903,78
Pesado R ²	7,83%
Wald chi ²	54,86
Acerto do modelo	87,01%

O primeiro modelo testado (Modelo 1) teve como inspiração o modelo desenvolvido por Costa (2012a) que utilizou a mesma base de dados. Contudo foram feitos alguns ajustes, como a não divisão dos resultados das variáveis explicativas por percentis, nem a condicionante relativamente à posse de dívidas por parte das famílias e a consideração apenas da base de dados original do ISFF de 2010. Estas diferenças são a razão pela qual os resultados não são semelhantes. Neste modelo foram consideradas 1.247 famílias. Apenas as variáveis rendimento, idade e deterioração das condições se mostraram

estatisticamente significativas para explicar o incumprimento das famílias, apesar do coeficiente da variável rendimento não ser expressivo.

Relativamente à influência e significância das variáveis explicativas no incumprimento das famílias, a variável riqueza líquida tem um impacto negativo no incumprimento como já era esperado, apesar de neste modelo não ser uma variável significativa para explicar o incumprimento, como ocorreu no estudo de Costa (2012a), onde também o seu coeficiente fora mais expressivo (superior a 1); o valor de dívida da família também não é estatisticamente significativo; o valor de rendimento apesar de significativo, o seu coeficiente não é expressivo; a idade do indivíduo de referência tem um impacto negativo na probabilidade de incumprimento em sintonia com o era esperado, apesar do seu coeficiente não ser tão expressivo quando comparado com o de Costa (2012a) (0,24); a escolaridade não se mostra significante, diferente do estudo de Costa (2012a), com um coeficiente menor que no modelo realizado pelo referido autor (0,9) para um nível de escolaridade superior; a situação profissional, que distinguiu apenas estar desempregado ou não, mostra que estar desempregado tem um impacto positivo no incumprimento, apesar de não significativo, não sendo o seu coeficiente tão relevante como o da análise de Costa (2012a) (0,76); por fim, a ocorrência de alterações desfavoráveis na situação financeira demonstra, tal como análise de Costa (2012a), ter um impacto positivo significativo no incumprimento, apesar do seu coeficiente ser inferior ao do referido autor (superior a 1).

Tabela 8 – Regressão probabilística para probabilidade de incumprimento das famílias
(Modelos 2, 3 e 4)

Regressão probabilística para probabilidade de incumprimento			
	Modelo 2	Modelo 3	Modelo 4
Riqueza líquida	-7,6E-07 (6,4E-07)	-1,5E-06 ** (7,4E-07)	-1,6E-06 ** (8,3E-07)
Dívida	5,2E-07 (1,0E-06)	-6,2E-08 (1,8E-06)	8,5E-07 (1,8E-06)
Rendimento	-7,7E-06 *** (3,0E-06)	2,0E-06 (4,2E-06)	3,0E-06 (4,2E-06)
Idade	-0,0083 * (0,0050)	0,0030 (0,0101)	0,0034 (0,0100)
Escolaridade	-0,0707 (0,1564)	-0,0124 (0,2669)	-0,0414 (0,2706)
Situação profissional	0,1516 (0,1181)		
Deterioração das condições	0,6427 *** (0,1161)	0,5549 *** (0,2119)	0,5153 ** (0,2124)
Número de crianças	0,1453 *** (0,0502)	0,2223 ** (0,0974)	0,2256 ** (0,0985)
Recusa de crédito		1,1414 *** (0,2108)	1,1229 *** (0,2132)
Duração do emprego atual		-0,0262 ** (0,0111)	-0,0226 ** (0,0113)
Propriedade de habitação			-0,3670 (0,2447)
Constante	-1,2265 *** (0,2803)	-1,7683 *** (0,5354)	-1,5833 *** (0,5382)
Número de observações	1.247	364	364
Log	-439,61	-122,92	-121,74
AIC	897,21	265,83	265,48
Pesado R ²	8,72%	24,56%	25,28%
Wald chi ²	57,99	65,81	65,17
Acerto do modelo	86,93%	85,16%	85,99%

No segundo modelo (Modelo 2), foi acrescentada a variável número de crianças do agregado, mantendo-se as mesmas observações que o modelo 1. Esta variável considerada por vários autores (Canner *et al.*, 1991; Zhang e

DeVaney,1999; Getter, 2003 e Duygan e Grant, 2009), mostra-se estatisticamente significativa e com influência positiva na probabilidade de incumprimento, com um coeficiente próximo da análise de Duygan e Grant (2009) (0,19) e menor que o que os conseguidos pelos restantes autores (com 0,29; 0,25 e 0,54). As restantes variáveis mantêm-se similares, em termos de significância e influência na probabilidade de incumprimento, ao modelo 1.

No terceiro modelo (Modelo 3), para além de excluir a variável situação profissional, foram acrescentadas as variáveis duração do emprego atual, considerada na análise de Zhang e DeVaney (1999) e a variável recusa de crédito considerada na análise de Godwin (1999). Este modelo contou com 364 observações onde ambas as variáveis incluídas se mostraram estatisticamente significativas, tal como as análises dos referidos autores faziam prever.

A variável duração do emprego atual incluída a partir do terceiro modelo, veio corroborar o estudo de Zhang e DeVaney (1999) ao ter um efeito negativo na probabilidade de incumprir com coeficiente inferior (0,0044) ao encontrado neste modelo. Contudo diverge ao se mostrar estatisticamente relevante, esta variável foi incluída em detrimento da situação profissional, pois para além de se mostrar significativa, está altamente correlacionada com a situação profissional. A ocorrência de recusa de crédito mostra ter um impacto positivo na probabilidade de incumprimento, o seu coeficiente é superior à análise de Godwin (1999) (0,59) que utilizou dados dum inquérito à famílias que data de 1989.

No quarto modelo (Modelo 4) foi adicionada a variável propriedade de habitação, incluída nas análises de Zhang e DeVaney (1999), Stavins (2000), Getter (2003) e Duygan e Grant (2009), mantendo-se a consideração de 364 observações, não se tendo verificado ser estatisticamente relevante para explicar o incumprimento das famílias, em semelhança à análise Stavins (2000). Constata-se ainda que o seu impacto é negativo à semelhança das análises dos referidos autores. O seu coeficiente é inferior às análises de Zhang e DeVaney

(1999), Getter (2003) e Duygan e Grant (2009) (0,61; 0,69 e 0,59), mas superior ao encontrado por Stavins (2000) (0,14).

Tabela 9 – Regressão probabilística para probabilidade de incumprimento das famílias (Modelo 5)

Regressão probabilística para probabilidade de incumprimento	
	Modelo 5
Riqueza líquida	-1,7E-06 * (9,2E-07)
Rendimento	3,1E-06 (4,1E-06)
Idade	0,0026 (0,0098)
Escolaridade	-0,0279 (0,2681)
Deterioração das condições	0,5211 ** (0,2133)
Número de crianças	0,2263 ** (0,0980)
Recusa de crédito	1,1145 *** (0,2128)
Duração do emprego atual	-0,0228 ** (0,0113)
Propriedade de habitação	-0,3248 (0,2317)
Constante	-1,5406 *** (0,5232)
Número de observações	364
Log	-121,87
AIC	263,73
Pesado R ²	25,20%
Wald chi ²	64,60
Acerto do modelo	85,99%

Por fim, o último modelo (Modelo 5) resultou da análise estatística do modelo anterior (Modelo 4), onde foi excluída a variável dívida, sendo este o modelo proposto.

A variável dívida denotou um efeito positivo na maioria dos modelos testados (Modelos 1, 2 e 4) à semelhança do estudo de Costa (2012a). No entanto, fora excluída no modelo final por não se mostrar significativa, piorar o ajuste do modelo (AIC) e ter sido rejeitada pelo teste de igualdade das médias.

A variável rendimento, apesar de indicar um efeito negativo e significativo nos dois primeiros modelos (Modelos 1 e 2), à semelhança dos modelos propostos por Godwin (1999), Stavins (2000), Duygan e Grant (2009), Alfaro e Gallardo (2012) e Costa (2012a), os últimos modelos (Modelos 3, 4 e 5) testados mostraram o efeito contrário, tal como a análise de Zhang e DeVaney (1999), não sendo porém significativa.

A variável idade também, nos primeiros dois modelos (Modelos 1 e 2), apresenta um efeito negativo e significativo, embora este efeito vá em sintonia com as análises de Canner *et al.* (1991), Godwin (1999), Stavins (2000), Getter (2003) e Costa (2012a), não se coaduna com as análises de Zhang e DeVaney (1999), Duygan e Grant (2009) e Alfaro e Gallardo (2012) e com os últimos modelos que foram testados (Modelos 3, 4 e 5).

As variáveis rendimento e idade apesar de não serem estatisticamente significantes a sua inclusão melhorou o ajustamento do modelo e daí terem sido incluídas no modelo final.

3.4.3 O Modelo Final

O modelo final proposto (Modelo 5), apresentado na Tabela 9, considerou para o ajustamento do modelo 364 famílias, onde 83,5% declaram não ter ocorrido qualquer atraso ou falha no cumprimento das suas obrigações e 16,5% declaram que incumpriram as suas responsabilidades no último ano (2009), embora o número de famílias não seja elevado e amostra não seja balanceada é de salientar que o incumprimento não é um evento frequente e que nem todas as famílias possuem responsabilidades o que leva a *missing values* relativamente à variável dependente considerada (o incumprimento ou mora de empréstimos).

A determinação do modelo derivou das variáveis selecionadas no quarto modelo (Modelo 4) testado, culminando em 10 passos e resultando na inclusão de 9 variáveis explicativas, como é evidente na Tabela 10.

Tabela 10 – Passos do modelo proposto

Passo	Variáveis	AIC	Log	<i>p-value</i>	Wald <i>chi</i> ²	Pseudo-R ²	Acerto do modelo
0	Y	1484	-741				
1	Y X ₉	487	-241	0	88,37	15,66%	84,07%
2	Y X ₉ X ₇	478	-236	0	95,34	17,52%	84,07%
3	Y X ₉ X ₇ X ₈	470	-231	0	99,90	19,40%	85,30%
4	Y X ₉ X ₇ X ₈ X ₆	325	-157	0	69,54	20,55%	85,62%
5	Y X ₉ X ₇ X ₈ X ₆ X ₁	258	-123	0	62,72	24,47%	85,44%
6	Y X ₉ X ₇ X ₈ X ₆ X ₁ X ₁₀	258	-122	0	62,84	25,05%	85,44%
7	Y X ₉ X ₇ X ₈ X ₆ X ₁ X ₁₀ X ₃	260	-122	0	64,20	25,18%	85,71%
8	Y X ₉ X ₇ X ₈ X ₆ X ₁ X ₁₀ X ₃ X ₂	262	-122	0	64,70	25,24%	85,71%
9	Y X ₉ X ₇ X ₈ X ₆ X ₁ X ₁₀ X ₃ X ₄	262	-122	0	64,36	25,20%	85,71%
10	Y X ₉ X ₇ X ₈ X ₆ X ₁ X ₁₀ X ₃ X ₄ X ₅	264	-122	0	64,60	25,20%	85,99%

A variável dívidas foi excluída do modelo por piorar o ajustamento do mesmo e por não ser estatisticamente diferente relativamente à variável dependente (*teste t*), por outro lado as variáveis idade e escolaridade apesar de

não melhorarem o ajustamento, a sua inclusão melhora o acerto do modelo, todas as restantes variáveis foram aceites no modelo por melhorarem o ajustamento (AIC) e os resultados gerais do modelo (*wald chi2*, pseudo-R² e acerto do modelo).

3.4.4 Avaliação do modelo proposto

3.4.4.1 Precisão do modelo

Para verificar a capacidade de precisão da previsão do modelo proposto foi apurada a aderência do modelo através do *teste de Hosmer e Lemeshow* (1980) e erros e acertos. No modelo proposto o resultado do *teste de Hosmer e Lemeshow* (1980), indica uma boa aderência do modelo entre os valores observados e os previstos, já que não se pode rejeitar a hipótese nula ($p\text{-value} > 0,05$)

Tabela 11 – Teste de Hosmer-Lemeshow

<i>Chi-square</i>	6,29
<i>p-value</i>	0,61

Para avaliar a capacidade de previsão do modelo proposto recorreu-se ainda à matriz de classificação, de forma a mensurar o número de famílias classificadas correta ou incorretamente pelo modelo proposto, conforme demonstra a Tabela 12, onde o modelo consegue prever corretamente o cumprimento e incumprimento de cerca de 85,99% das famílias consideradas, que nesta amostra representam 313 famílias. Por outro lado, o modelo prevê erradamente 12 famílias como sendo incumpridoras, quando na verdade estas declararam ter cumprido os seus compromissos. Em sentido inverso, 39 famílias foram classificadas como cumpridoras, quando na verdade declaram ter incumprido as suas obrigações.

Tabela 12 – Matriz de classificação do modelo proposto

Estimado	Observado			Classificações corretas
	Incumprimento	Cumprimento	Total	
Incumprimento	21	12	33	85,99%
Cumprimento	39	292	331	
Total	60	304	364	

3.4.4.2 Medidas de ajustamento do modelo

Para mensurar o ajustamento foram usadas algumas medidas, como o critério de informação de Akaike, o *log* de verossimilhança, heterocedasticidade do modelo e o valor do pseudo R^2 .

O Critério de Informação de *Akaike* (AIC), Akaike (1973), foi a medida usada para determinar a qualidade do ajustamento do modelo estimado. Este critério determina a distância entre o modelo estimado e o observado, assim o modelo escolhido será o que minimizar o valor de AIC, ou seja, $AIC^K < AIC^{K-1}$. Assim como se observa na Tabela 10 no primeiro passo para estimação do modelo o valor era 1.484, e no modelo final apenas 264, no que diz respeito ao AIC. Também ao se examinar os modelos estimados na calibragem do modelo (Tabelas 7, 8 e 9) denota-se que no primeiro modelo testado (Modelo 1) tem um valor no teste AIC de cerca de 904 e culmina no último modelo (Modelo 5) num valor de 264.

O *log* de verossimilhança, tal como o AIC, é um indicador de ajuste do modelo, e tal como AIC quanto menor o seu valor melhor o ajuste do modelo. Como demonstra, na Tabela 10, o primeiro passo do modelo tinha o valor de -741 e o modelo final apenas -122, no que diz respeito ao *log* de verossimilhança. Também ao se contemplar os modelos estimados na calibragem do modelo (Tabelas 7, 8 e 9) denota-se que no primeiro modelo testado (Modelo 1) tem um valor de -444 e culmina no último modelo (Modelo 5) num valor de -122.

Numa regressão linear múltipla um dos pressupostos é que as variâncias dos erros não sejam função das variáveis explicativas, o mesmo será dizer que não haja correlação entre o erro (ε_i) e as variáveis explicativas (x_i). O teste de Breusch-Pagan (1979) testa se a variância do erro é constante ou se a variância do erro é função dos valores das variáveis explicativas.

O modelo proposto apresenta os seguintes valores quando estimado o teste de Breusch-Pagan:

Tabela 13 – Teste de Breusch-Pagan

<i>Chi-square</i>	98,71
<i>p-value</i>	0,00

A Tabela 13 mostra que o modelo proposto sofre de heterocedasticidade problema típico quando os dados são *cross-section* como é o caso deste estudo. Para corrigir este problema recorreu-se à aplicação “*robust standard errors*”, que assume que os erros são distribuídos de forma independente.

Por fim, o pseudo R^2 , definido por McFadden (1974), é também um indicador de ajuste geral do modelo. Como apresenta a Tabela 10 no primeiro modelo o valor era de 15,66% e no modelo final 25,2%, no que diz respeito ao pseudo R^2 . A mesma melhoria aconteceu nos modelos que serviram para calibrar (Tabelas 7, 8 e 9) o modelo onde o primeiro modelo testado (Modelo 1) apenas possuía 7,83% valor menor que o do último modelo (Modelo 5).

3.4.4.3 Validação do modelo

Como demonstrado anteriormente o modelo proposto classificou corretamente 85,99% das famílias. Contudo não se pode considerar que ele funcionará se mudar a realidade, assim este modelo é considerado um bom modelo, tendo em conta a amostra em que se sustenta. São necessários testes de

validação, principalmente porque se pretende que seja um modelo de previsão de incumprimento (Hosmer e Lemeshow, 2013).

Para validar o modelo recorreu-se ao método *jackknife*, de Lachenbruch (1967), que permite a utilização de toda a amostra no desenvolvimento do modelo, tendo por base o princípio *leave-one-out* será separada uma observação da amostra original estimando-se um novo modelo sem a respetiva observação selecionada, este processo é então repetido para todas as observações.

A Tabela 14 apresenta a matriz de classificação obtida pela utilização deste método, que se manteve igual à Tabela 12, de onde que se conclui a validade do modelo proposto.

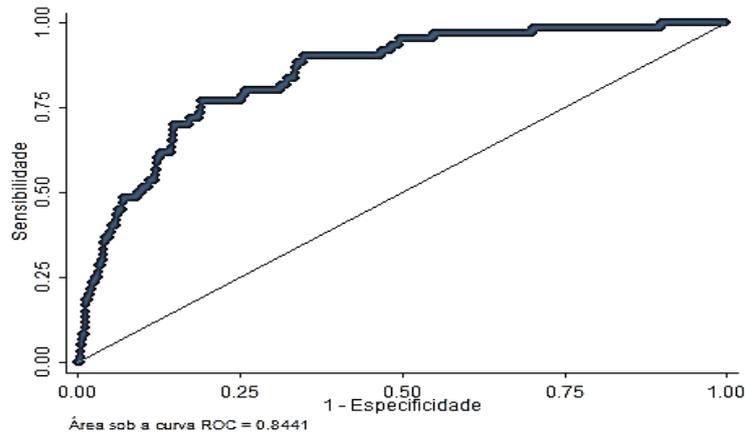
Tabela 14 – Matriz de classificação resultado da validação do modelo

Estimado	Observado			Classificações corretas
	Incumprimento	Cumprimento	Total	
Incumprimento	21	12	33	85,99%
Cumprimento	39	292	331	
Total	60	304	364	

A curva ROC (*Receiver Operating Characteristic*) é outro processo usado para avaliar o modelo, baseado nos conceitos de sensibilidade e especificidade, onde a sensibilidade é proporção de acerto, no caso de ocorrência do evento tendo ele ocorrido, e onde a especificidade é proporção de acerto no caso da não ocorrência do evento, tendo ele não ocorrido.

A curva ROC do modelo proposto, Gráfico 5, indica uma área sob a curva ROC de 0,8441, logo segundo as regras sugeridas por Hosmer e Lemeshow (2013) sugere uma excelente discriminação.

Gráfico 5 – Curva ROC



3.5 Classes de risco

Finda a análise estatística e validado um modelo capaz de prever o incumprimento das famílias portuguesas cabe determinar classes de risco de forma a puderem ser geradas notações de risco. A probabilidade de incumprimento estimada no modelo exposto assume valor de 0 a 1, logo existirá uma infinidade de partições possíveis, a divisão feita se designará de classe de risco. Para determinar o número de classes recorreu-se ao método de Calinski e Harabasz (CH) (1974), dado por:

$$CH(k) = \frac{\sum_{i=1}^k \frac{n_i(\bar{y}_i - \bar{y})^2}{(k-1)}}{\sum_{i=1}^k \sum_{j=1}^{n_i} \frac{n_i(y_{ij} - \bar{y}_i)^2}{(n-k)}} \quad (3.2)$$

onde k é o número de classes, n o número de observações, i a identificação da classe e y_{ij} a observação j estimada na classe de risco i .

O método de partição usado foi o método de *Ward* que minimiza a distância (euclidiana) entre a probabilidade de incumprimento média da classe (\bar{y}_c) e a observada (y_i).

Todo o processo teve por base as probabilidades previstas pelo modelo proposto nesta exposição, pelo que uma amostra diferente poderá sugerir um número distinto de classes de risco. Através da aplicação dos métodos enunciados anteriormente obteve-se 15 classes, contudo foram acrescentadas uma classe inicial e uma final de modo a que probabilidade de incumprimento reflita um valor contínuo de 0% a 100%, considerou-se ainda o valor máximo de cada classe como o limite mínimo da classe seguinte.

Definido o número de classes ótimo, a Tabela 15 reflete os intervalos de classificação estimados para cada classe.

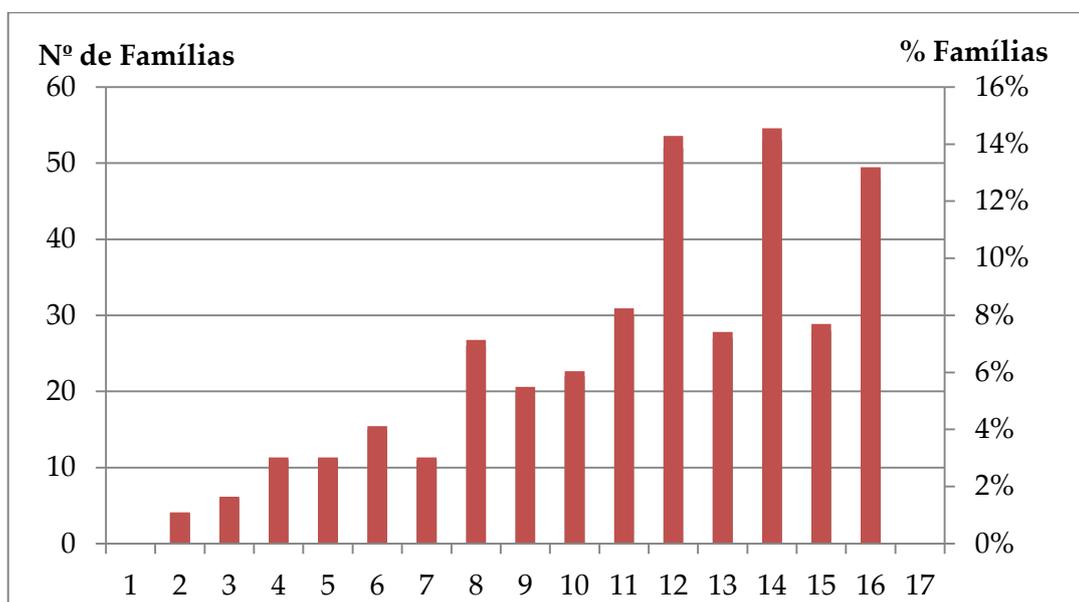
Tabela 15 – Intervalos de classificação para probabilidade de incumprimento

Classes	Intervalos de classificação (Probabilidade de Y=1)	
17	0,000%	0,002%
16	0,002%	1,931%
15	1,931%	3,004%
14	3,004%	5,466%
13	5,466%	7,167%
12	7,167%	10,754%
11	10,754%	14,501%
10	14,501%	17,418%
9	17,418%	22,337%
8	22,337%	29,499%
7	29,499%	38,217%
6	38,217%	50,407%
5	50,407%	59,810%
4	59,810%	65,844%
3	65,844%	74,089%
2	74,089%	87,725%
1	87,725%	100,000%

Assim se pode concluir que quanto maior a classe, menor a probabilidade de incumprimento da família. Esta ordenação guiou-se pelas notações já desenvolvidas para particulares, como a FICO score e a VantageScore, embora com uma escala numérica esta não tem a mesma extensão das notações enunciadas.

Posto isto, para a amostra deste estudo, verifica-se a seguinte distribuição de observações por classe, conforme mostra o Gráfico 6, que permite concluir que mais de 50% das famílias consideradas têm uma probabilidade de incumprimento inferior a 10,754%, que corresponde a estar numa classe igual ou superior à classe 12. Por outro lado, cerca de 3% das famílias têm um valor elevado de probabilidade de incumprimento superior a 65,844%, que se traduz em estar na classe 3 ou abaixo.

Gráfico 6 – Distribuição da amostra pelas classes de risco



3.6 Perda esperada

O risco de crédito, como enunciado no Capítulo 2, é mensurado pela perda esperada e pela perda não esperada. Para determinar a perda esperada é necessário para além da probabilidade de incumprimento é ainda indispensável determinar a exposição em caso de incumprimento e a perda em caso de incumprimento.

Para determinar a exposição a incumprimento da família foram assumidos alguns pressupostos:

- i) O valor de pagamentos mensais de empréstimos declarados no momento da realização do inquérito se manterá constante durante os próximos 12 meses;
- ii) O valor em dívida, por parte da família, não sofrerá aumentos ao longo de 12 meses;
- iii) Famílias que não declarem valor de pagamentos mensais, mas que detenham valor de dívidas assumem EAD igual a 1, ou seja assume-se que o valor em dívida se manterá nos próximos 12 meses;
- iv) Famílias que declarem valor de pagamentos mensais que acumulados são superiores ao valor de dívidas no espaço de 12 meses assume-se um EAD igual a zero, ou seja, não existirá qualquer valor exposto a incumprimento.

Assumidos os pressupostos referidos a exposição a incumprimento num período de 12 meses, será dado pela fórmula:

$$EAD_i = \frac{D_i - R_i}{D_i} \times 100 \quad (3.3)$$

$$VEAD_i = EAD_i \times D_i \quad (3.4)$$

onde D_i corresponde ao valor das total das dívidas da família i no momento de realização do inquérito e R_i corresponde ao valor que se espera que família reembolse no prazo de 12 meses.

Considerando apenas as 364 famílias determinantes para modelar o modelo proposto e para as quais se conseguiu determinar uma probabilidade de

incumprimento, pode-se concluir que em média estas famílias apresentam cerca de 73,59% em termos de EAD_i . Em termos de valor ($VEAD_i$) em média cada família, passados 12 meses, terá um valor em dívida de aproximadamente 49 mil euros.

Para determinar a perda em caso de incumprimento das famílias utilizou-se a mesma fórmula de Albacete e Lindner (2013), num estudo que o utilizou dados do mesmo inquérito (HFCS de 2010) mas com os dados da Áustria, dada por:

$$LGD_i = \frac{PD_i \times (D_i - RB_i) \times NRL_i}{D_i} \quad (3.5)$$

onde PD_i é a probabilidade de incumprimento das famílias determinada pelo modelo proposto para a família i , D_i corresponde ao valor de dívidas da família, RB_i é o valor de riqueza bruta declarado pela família e NRL_i é um indicador que assume valor 1 caso a riqueza líquida da família seja negativa e zero caso contrário.

Constatou-se que das 364 famílias cerca de 45% possui um valor de riqueza líquida negativa, sendo para estas determinado um valor de perda em caso de incumprimento, que ronda os 13,4%.

Em conclusão, possuindo a probabilidade de incumprimento (PD) determinada no modelo proposto neste estudo, a exposição a incumprimento dentro de 12 meses (EAD) e a perda esperada em caso de incumprimento (LGD) tem-se todos os componentes necessários para a determinação da perda esperada (EL), dada pela fórmula:

$$EL_i = PD_i \times EAD_i \times LGD_i \quad (3.6)$$

$$VEL_i = EL_i \times D_i \quad (3.7)$$

Considerando as famílias com riqueza líquida negativa ($LGD_i \neq 0$) em termos médios percentuais tem-se cerca de 4,2% de EL_i , que em valor (VEL_i) representa cerca de 3 mil euros, com um valor máximo calculado de cerca de 44 mil euros-

4 Conclusões

Num momento em que ainda se considera de recuperação da crise financeira de 2007 e onde já muito fora mencionado como estando na sua origem, a medição da probabilidade de incumprir assume grande destaque, não só por ter sido apontada como uma das causas que estiveram na origem da crise, mas também justificada pelo facto do sistema financeiro se ter tornado mais cauteloso e mais regulado do que era até 2007.

O presente Trabalho Final de Mestrado pretendeu contribuir para reflexão nesta matéria, desencadeando um modelo de medição de probabilidade de incumprimento, usando dados do ISFF de 2010 realizado em Portugal, e a respetiva definição de classes de risco. Acessoriamente também foi determinada a perda esperada, principal componente do risco de crédito.

O interesse deste trabalho resulta do seu contributo para: i) esclarecer conceitos como o crédito, risco de crédito e notação de risco; ii) a reunião de revisão teórica e empírica necessária ao desenvolvimento de um modelo de previsão de incumprimento de particulares ou famílias; iii) a distinção dos tipos de agências que se dedicam ao desenvolvimento e emissão de notações de risco; iv) o enquadramento do crédito a particulares em Portugal, onde se destaca a variação de crédito a particulares, a evolução do rácio de crédito de vencido e do rácio dívida/rendimento; v) um breve enquadramento das agências de notação existentes em Portugal, na Europa e a nível mundial; e por fim vi) a determinação de um modelo de medição e notação de incumprimento das famílias portuguesas.

O modelo proposto aplicado ao Inquérito à Situação Financeiras das Famílias (ISFF) guiou-se por estudos similares já desenvolvidos, destacando-se o de Godwin (1999), Alfaro e Gallardo (2012), Costa (2012a) e Albacete e Lindner (2013), embora em nenhum dos estudos enunciados tenha sido proposta uma

escala de forma a ser possível atribuir uma notação de risco, como neste trabalho.

No que diz respeito aos resultados do modelo constatou-se que a declaração de piores condições, comparativamente aos últimos 3 anos, sejam estas de trabalho, rendimento, riqueza ou emprego, são um indício muito forte relativamente à ocorrência de incumprimento da família. Também o número de crianças do agregado revela que quando maior o número de crianças maior será a probabilidade do agregado incumprir. Por fim, cabe salientar que também a ocorrência de recusa de crédito nos últimos 12 meses pesa positivamente para um aumento da probabilidade de incumprimento. No sentido inverso, se o nível de escolaridade declarado pelo representante da família no inquérito for superior, menor será à partida a probabilidade de incumprimento da família e o mesmo se sucede no caso de o agregado seja detentor de habitação própria. No que respeita ao valor de riqueza líquida comprovou que à medida que este for maior menor será a probabilidade de incumprimento, contudo o seu peso não se mostrou significativo. A probabilidade de incumprimento diminui ainda quanto maior o número de anos empregado na empresa atual declarados pelo representante da família.

O modelo proposto permitiu prever com um nível de acerto elevado (85,99%) a probabilidade de uma família incumprir com base nas variáveis utilizadas para o modelar. Ao permitir prever o incumprimento poderá não só servir de guia para justificar a atribuição de crédito, no caso de ser utilizado por uma instituição de crédito, mas também servir de base para a atribuição de uma notação de risco de uma qualquer agência de notação.

A definição de classes de notação de risco, para a amostra considerada no modelo proposto, é o último aspeto deste estudo que permitiu assim cumprir o objetivo definido que era a proposta de um modelo de notação de risco. Como tal, foram propostas 17 classes de risco, onde pertencer à classe 17 significa menor probabilidade de incumprimento e pertencer à classe 1 significa ter uma

grande probabilidade de incumprimento ($PD > 87,725\%$). Numa breve análise da distribuição da amostra (364 famílias) pelas classes de notação definidas concluiu-se que mais de 50% das famílias tem uma baixa probabilidade de incumprir as suas obrigações ($PD < 10,754\%$), ou seja, o mesmo será dizer que 50% tem uma notação de risco igual ou superior a 12.

5 Limitações do estudo

Como limitações deste estudo destaca-se a inviabilidade de interconexão de dados, a interconexão de dados da Central de Responsabilidades de Crédito (CRC) sob a alçada do BdP, da lista de devedores da Autoridade Tributária e Aduaneira e da lista de devedores à Segurança Social. Além da não viabilidade de reunir fontes diferentes, estas não seriam suficientes para definir um modelo de previsão de incumprimento, já que estas são maioritariamente compostas por dados financeiros.²²

Outra limitação encontrada foi a base de dados ser composta apenas por dados recolhidos do ISFF de 2010, limitação que não foi possível ultrapassar já que esta foi a primeira edição com este formato. Contudo, em 2013 foi realizada uma segunda edição mas os dados ainda não se encontram disponíveis. A junção de mais observações tornaria com certeza o modelo desenvolvido mais robusto.

²² Apesar do FICO score e do VantageScore utilizar apenas dados relativos a crédito, os modelos de notação revistos na literatura não contêm somente este tipo de variáveis, como se denota nos anexos 1 e 2.

6 Pesquisa futura e sugestões

A notação de risco para famílias, proposta neste trabalho, tem ainda um longo caminho a percorrer, já que em Portugal a única agência de notação ainda não se dedica à notação de particulares, o que pode ser justificado pela falta de informação deste tipo de devedores, mas também por motivos culturais. Assim antes de mais é necessário uma mudança cultural que para além duma aposta na formação financeira, necessária para um melhor entendimento das vantagens que este tipo de notação poderá trazer, para devedores, particulares, e para as instituições que concedem crédito.

Como mencionado por Ross *et al.* (2003) não existem fórmulas mágicas para medir a probabilidade de um devedor incumprir, como qualquer modelo de notação de risco baseiam-se em regras (Huang *et al.*, 2007) que esperam que originem um modelo capaz de prever corretamente o risco de determinado devedor, e como tal também o modelo aqui proposto se espera que tenha um contributo positivo na busca dessa fórmula mágica.

No que diz respeito a sugestões seria pertinente serem desenvolvidos modelos de previsão de incumprimento usando dados de outros países da EU já que o ISFF é de âmbito europeu, apesar de algumas questões não constarem do inquérito em todos os países. A questão relativa a atraso ou incumprimento de empréstimos, que serviu para definir a variável dependente deste estudo, conforme explicito no catálogo das variáveis do ISFF de 2010, foi apenas incluída nos inquéritos realizados em Portugal, Espanha e Luxemburgo (HFCN, 2012b). De qualquer forma, comparar o modelo entre vários países seria sempre enriquecedor. Por outro lado, a junção das bases de dados de diferentes países para modelar a probabilidade de incumprimento, além aumentar o número de observações seria interessante verificar a divergência ou consistência dos resultados da modelação individual de cada país. Também a junção de dados

de várias edições do ISFF enriquecerá certamente a modelação da probabilidade de incumprimento das famílias.

Referências bibliográficas

Abdou, H., e Pointon, J. 2011. Credit scoring, statistical techniques and evaluation criteria: a review of the literature, *Intelligent Systems in Accounting, Finance & Management*: 59-88.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In B. N. P. e. F. Csaki (Ed.), *Second International Symposium on Information Theory*: 267–281. Budapest.

Akerlof, G. 1970. The market for lemons, Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 83 (3): 488-500.

Albacete, N., e Lindner, P. 2013. Household Vulnerability in Austria – A Microeconomic Analysis Based on the Household Finance and Consumption Survey, *Financial Stability Report (junho 2013)*: 57-73: OeNB.

Alfaro, R., e Gallardo, N. 2012. The Determinants of Household Debt Default. *Revista de Analisis Económico*, 27 (1): 55-70.

Alves, N., e Ribeiro, N. 2011. Modelação do incumprimento dos particulares, *Relatório de Estabilidade Financeira (Novembro 2011)*: 191-201. Lisboa: BdP.

Anderson, R. 2007. *The credit scoring toolkit: Theory and practice for retail credit risk management and automation*: Oxford University Press Inc.

Avery, R., Bostic, R., Calem, P. e Canner, G. 2000. Credit scoring: statistical issues and evidence from Credit-bureau files. *Real Estate Economics*, 28(3): 523-547.

Baker, H., e Mansi, S. 2002. Assessing Credit Rating Agencies by Bond Issuers and Institutional Investors. *Journal of Business Finance and Accounting*, 29(9): 1367-1398.

Banasik, J., e Crook, J. 2007. Reject inference, augmentation, and sample selection. *European Journal of Operational Research*, 183: 1582–1594.

Banco de Portugal. 2011. Relatório do inquérito à literacia financeira da população portuguesa 2010. Lisboa.

Banco de Portugal. 2014a. Relatório de Estabilidade Financeira Maio 2014. Lisboa.

Banco de Portugal. 2014b. Projeções para a Economia Portuguesa: 2014-2016. *Boletim Económico*: 55-59.

- Bernanke, B., Gertler, M. e Gilchrist, S. 1996. *The Financial Accelerator and the Flight to Quality*: 1-15: The MIT Press.
- Bhatia, A. 2002. *Sovereign Credit Ratings Methodology: An Evaluation*: 60: International Monetary Fund, IMF Working Papers.
- Boot, A., Milbourn, T., e Schmeits, A. 2003. *Credit ratings as coordination mechanisms*. St. Louis.: Washington University.
- Breusch, T. e Pagan, A. 1979. Simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 45 (5): 1287–1294.
- Brill, J. 1998. The importance of credit scoring models in improving cash flow and collections. *Business Credit*, 100 (1).
- Burns, B. 2010. *Keeping Score on Credit Scores: An Overview of Credit Scores, Credit Reports and Their Impact on Consumers*. Stamford: VantageScore Solutions.
- Butler, A., e Cornaggia, K. 2012. Rating through the relationship: Soft information and credit ratings: 35.
- Calinski, T., e Harabasz, J. 1974. A dendrite method for cluster analysis. *Communications in Statistics* (3): 1–27.
- Canner, G., Gabriel, S. e Woolley J. 1991. Race, Default Risk and Mortgage Lending: A Study of the FHA and Conventional Loan Markets. *Southern Economic Journal*, 58(1): 249-262.
- Caouette, J., Altman, E., Narayanan, P. e Nimmo, R. 2008. The Rating Agencies; Introduction to Credit Risk Models; Testing and Implementation of Credit Risk Models, *Managing credit risk: the great challenge for global financial markets*, 2 ed.: 81-101, 127-137 e 237-250. New Jersey: John Wiley & Sons, Inc.
- Carvalho, P. 2009. *Fundamentos da Gestão de Crédito*. Lisboa.
- Cordeiro, P. 2007. Avaliar risco de crédito com pouca informação. *Diário de Notícia*. Disponível em http://www.dn.pt/inicio/interior.aspx?content_id=650962 consultado a 26.08.2014.
- Chuang, C., e Lin, R. 2009. Constructing a reassigning credit scoring model. *Expert Systems with Applications*, 36 (2): 1685–1694.

Costa, S. 2012a. Probabilidade de incumprimento das famílias: uma análise com base nos resultados do ISFF, *Relatório de Estabilidade Financeira (Novembro 2012)*: 101-115. Lisboa: BdP.

Costa, S., e Farinha, L. 2012b. Inquérito a Situação Financeira das Famílias: metodologia e principais resultados: BdP.

Costa, S., e Farinha, L. 2012c. O endividamento das famílias: uma análise microeconómica com base nos resultados do inquérito a situação financeira das famílias., *Relatório de Estabilidade Financeira (Maio 2012)*: 137-164. Lisboa: BdP.

Davis, R., Edelman, D. e Gammerman, A. 1992. Machine-learning algorithms for credit-card applications. *IMA Journal of Management Mathematics*, 4 (1): 43-51.

DefaultRisk.com. 2014. Credit Rating Agencies - globally. Disponível em http://www.defaultrisk.com/rating_agencies.htm consultado a 15.05.2014

Dimitriu, M., Avramescu, E. e Caracota, R. 2010. Credit scoring for individuals. *Academic Journal*, 13 (2): 361-377.

Durand, D. 1941. *Risk Elements in Consumer Instalment Financing*. New York: National Bureau of Economic Research.

Duygan, B., e Grant, C. 2009. Household Debt Repayment Behaviour: what role do institutions play?, *Economic Policy*: 107-140: Federal Reserve Bank of Boston.

Eisenbeis, R. 1977. Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics. *The Journal of Finance*, 32 (3): 875-900.

Ellis, D. 1998. Different Sides of the Same Story: Investors and Issuers Views of Rating Agencies. *Journal of Fixed Income*, 7 (4): 35-45.

ESMA. 2013. CRAs' Market share calculation according to Article 8d of the CRA Regulation. Disponível em http://www.esma.europa.eu/system/files/esma_cra_market_share_calculation.pdf consultado a 13.05.2014.

ESMA. 2014a. ESMA approves EuroRating as a credit rating agency. Disponível em http://www.esma.europa.eu/system/files/2014-477_esma_approves_eurorating_as_a_credit_rating_agency.pdf consultado a 15.05.2014.

ESMA. 2014b. List of registered and certified CRAs. Disponível em <http://www.esma.europa.eu/page/List-registered-and-certified-CRAs> consultado a 13.05.2014

Espin-García, O., e Rodríguez-Caballero. C. 2013. Metodología para un scoring de clientes sin referencias crediticias. *Cuadernos de Economía*, 32 (59): 137-161.

Farinha, L., e Lacerda, A. 2010. Incumprimento no crédito aos particulares: qual e o papel do perfil de crédito dos devedores?, *Relatório de Estabilidade Financeira (Novembro 2010)*: 145-165. Lisboa: BdP.

Feio, D., e Carneiro, B. 2012. *O Poder das Agências - o que fazem, como falham e o que mudar nas agências de rating* (2 ed.).

FICO. 2005. Understanding your FICO Score. Disponível em http://www.myfico.com/downloads/files/myfico_uyfs_booklet.pdf consultado a 08.05.2014.

FICO. 2014a. Our History. Disponível em <http://www.fico.com/en/about-us/history/> consultado a 06.09.2014

FICO. 2014b. FICO Score – Measure consumer risk with the world's leading credit risk score. Disponível em <http://www.fico.com/en/products/fico-score/> consultado a 06.09.2014.

Fisher, R. 1936. The use of multiple measurements in taxonomic problems, Vol. 7: 179–188: *Annals of Eugenics*.

Frade, C. 2007. *A regulação do sobreendividamento*. Universidade de Coimbra, Coimbra.

Gaspar, C. 2014. Risco de Crédito - A importância da gestão de carteiras de crédito. *Inforbanca* (100): 41-43.

Gavras, P. 2012. Ratings Game. *Finance & Development*, 49 (1): 34-37.

Getter, D. 2003. Contributing to the Delinquency of Borrowers. *The Journal of Consumer Affairs*, 37 (1): 86-100.

Gitman, L. 2005. Risco e Retorno, *Princípios de Administração Financeira*, 10 ed.: 182-187. São Paulo.

Godwin, D. 1999. Predictors of Households' Debt Repayment Difficulties. *Financial Counseling and Planning*, 10 (1): 67-78.

- Gonçalves, E., Gouvêa, M. e Mantovani, D. 2013. Análise de risco de crédito com o uso de regressão logística, *Revista Contemporânea de Contabilidade*, Vol. 10: 139-160.
- Graça, J. 2012. Agências de notação de risco, *Cidepre*, Vol. 14: 1-25. Coimbra.
- Greene, W. 1992. A Statistical Model for Credit Scoring: 92-29. Nova Iorque: New York University.
- Hand, D., e Henley, W. 1997. Statistical Classification Methods in Consumer Credit Scoring: a Review. *Journal of the Royal Statistical Society*, 160 (3): 523-541.
- Harrell, F., e Lee, K. 1985. A comparison of the discrimination of discriminant analysis and logistic regression. In P. K. Sen (Ed.), *Biostatistics: Statistics in biomedical, public health, and environmental sciences*. Amsterdam.
- Haykin, S. 2001. Introdução, *Redes Neurais - Princípios e prática*, 2 ed.: 27-49: Prentice Hall Inc.
- Household Finance and Consumption Network. 2012a. Doc.UDB2 - HFCS Core Variables Catalogue: BCE.
- Household Finance and Consumption Network. 2012b. Doc.UDB4 - HFCS Non-core Variables Catalogue: BCE.
- Homer, S., e Sylla, R. 2005. Prehistoric and primitive credit and interest, *A History of Interest Rates*, 4 ed.: 17-31. New Jersey: John Wiley & Sons, Inc.
- Hosmer, D., e Lemeshow, J. 2013. Assessing the Fit of the Model, *Applied Logistic Regression*, 3 ed.: 153-222. New Jersey: John Wiley & Sons.
- Hosmer, J., e Lemeshow, S. 1980. Goodness-of-fit tests for the multiple logistic regression model: 1043–1069: *Communications in Statistics*
- Hovakimian, A., Kayhan A. e Titman, S. 2009. Credit Rating Targets: 1-50.
- Huang, C., Chen, M. e Wang, C. 2007. Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications*, 33: 847–856.
- IDC, F. I. 2013. Top 100 Companies in FinTec. Disponível em <http://www.americanbanker.com/fintech100/rankings/> consultado a 10.05.2014: American Banker e Bank Technology News.

IMF. 2010. The uses and abuses of sovereign credit ratings, *Global financial stability report: sovereigns, funding, and systemic liquidity*: 85-122.

Investe, P. 2009. Agências de 'rating' - Entidades cujo negócio é mesmo avaliar o risco dos outros. Disponível em <https://www.deco.proteste.pt/investe/agencias-de-rating-s1705681.htm#UE> consultado a 06.09.2014.

Jappelli, T., e Pagano, M. 2000. Information sharing in credit markets: A Survey. Itália: Centre for Studies in Economics and Finance, University of Naples.

Johnson, R. 2004. Legal, Social, and Economic Issues in Implementing Scoring in then United States, *Readings in Credit Scoring: recent developments, advances, and aims*: 5-15: Oxford University Press.

Johnson, R., e Wichern, D. 2007. Discrimination and classification, *Applied multivariate statistical analysis*, 6 ed.: 575-649. New Jersey: Prentice-Hall.

Jorion, P. 2007. The need for risk management, Lessons from financial disasters e Credit risk management, *Value at Risk, the new benchmark for managing financial risk*, 3 ed.: 1-49 e 451-490. Nova Iorque: McGraw-Hill.

Júlio, F. 2013. *Aplicação de Modelos de Credit Scoring na Gestão do Risco do Crédito no Sector Bancário Angolano Caso de Estudo: Bpc e Banco Sol*. Universidade Autónoma de Lisboa, Lisboa.

Kisgen, D. 2006. Credit Ratings and Capital Structure. *The Journal of Finance*, 61 (3): 1035-1072

Lachenbruch, P. 1967. An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis. *Biometrics*, 23: 639-645.

Lee, T., Chiu, C., Lu, C. e Chen, I. 2002. Credit scoring using the hybrid neural discriminant technique. *Expert Systems with Applications*, 23: 245-254.

Lee, T., e Chen, I. 2005. A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks and Multivariate Adaptive Regression Splines. *Expert Systems with Applications*, 28: 743-752.

Lehmann, B. 2003. Is It Worth the While? The Relevance of Qualitative Information in Credit Rating: 1-25: Universitaet Konstanz.

Lyons, A., Rachlis, M. e Scherpf, E. 2007. What's in a Score? Differences in Consumers' Credit Knowledge Using OLS and Quantile Regressions. *Journal of Consumer Affairs*, 41 (2): 223-249.

- MacCrimmon, K., Stanbury, W. e Wehrung, D. 1986. *Taking risks : the management of uncertainty*. Nova Iorque.
- Marques, M., e Frade, C. 2003. Uma sociedade aberta ao crédito. In Almedina (Ed.), *Sub Judice 24 - O Estado do Direito do Consumidor*: 27-34.
- McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*: 105–142.
- Medina, R., e Selva, M. 2012. Análisis del Credit Scoring, *RAE: Revista De Administração De Empresas*, Vol. 53: 303-315. São Paulo.
- Mester, L. 1997. What's the Point of Credit Scoring?, *Business Review*: 3-16: Federal Reserve Bank Of Philadelphia.
- Min, J., e Lee, Y. 2008. A practical approach to credit scoring. *Expert Systems with Applications*, 35: 1762–1770.
- Mishkin, F. 2004. An Economic Analysis of Financial Structure, *The Economics of Money, Banking and Financial Markets*, 7 ed.: 167-200.
- Moreira, F. 2011. O consumo e o crédito na sociedade contemporânea, Vol. 19: 91-114: Departamento de Economia, Gestão e Ciências Sociais – Centro Regional das Beiras da Universidade Católica Portuguesa.
- NACM. 2011. Who is the Customer? Revisiting the Five Cs of Credit, *Principles of Business Credit*, 6 ed.: 50-53. Colombia.
- Obrová, V. 2012. *Construction and application of scoring models*. Paper presented at the Proceedings of 30th International Conference Mathematical Methods in Economics.
- OCC. 2001. Rating Credit Risk Comptroller's Handbook. 1-67.
- Ong, C., Huang, J. e Tzeng, G. 2005. Building credit scoring models using genetic programming. *Expert Systems with Applications*, 29 (1): 41-47.
- Paiva, C. 1997. *Administração do risco de crédito*. Rio de Janeiro: Qualitymark Editora.
- Paliwal, M., e Kumar, U. 2009. Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36: 2-17.
- Pedro, C. 2012. Prefácio, *Quem atirou Portugal para o lixo? Agências de rating: o que são, como funcionam*, 2 ed.: Actual Editora.

Pereira, V. 2009. *Seleção adversa na concessão de Crédito a particulares em Portugal* Instituto Superior de Ciências do Trabalho e da Empresa, Lisboa.

Pereira, V. 2011. Seleção adversa no crédito a particulares em Portugal - Papel dos «bureaus» de crédito privados, *Revista Portuguesa e Brasileira de Gestão*: 14-22.

Pianesi, L. 2011. Le agenzie di rating tra privatizzazione di funzioni pubbliche e opinioni private "geneticamente modificate" *Rivista trimestrale di diritto pubblico*: 179-213.

Reichert, A., Cho, C. e Wagner, G. 1983. An examination of the conceptual issues involved in developing credit-scoring models. *Journal of Business and Economic Statistics*, 1 (2): 101-114.

Riestra, A. 2002. Credit bureaus in today's credit markets, *European Credit Research Institute*: 1-34.

Ross, S., Westerfield, R. e Jordan, B. 2003. *Fundamentals of Corporate Finance* (6 ed.). Boston: McGraw-Hill.

SaeR. 2014. ARC Ratings apresenta-se como alternativa no mercado mundial da notação de crédito. Disponível em <http://www.saer.pt/?lop=conteudo&op=f7177163c833dff4b38fc8d2872f1ec6&id=63538fe6ef330c13a05a3ed7e599d5f7> consultado a 13.05.2014.

Santos, J. 2003. *Análise de crédito – empresas e pessoas físicas* (2 ed.). São Paulo.

Saunders, A., e Allen, L. 2010. Setting the stage for financial meltdown; The three phases of the credit crisis; Other credit risk models, *Credit Risk Measurement In and Out of the Financial Crisis*, 3 ed.: 3-44 e 117-131. New Jersey: John Wiley & Sons, Inc.

Saunders, A., e Cornett, M. 2008. Credit Risk: Individual Loan Risk, *Financial institutions management: a risk management approach*, 6 ed.: 313-320: McGraw-Hill/Irwin.

Schreiner, M. 2000. Credit scoring for microfinance: Can it work? *Journal of Microfinance*, 2 (2): 105-118.

Schreiner, M. 2003. Scoring: the next breakthrough un microcredit? *Occasional Paper*, 7.

Schreiner, M. 2004. Benefits and Pitfalls of Statistical Credit Scoring for Microfinance. *Savings and Development*, 28 (1): 63–86.

Schreiner, M. 2005. Can Scoring Help Attract Profit-Minded Investors to Microcredit?, *Financial Sector Development Conference*. Frankfurt Germany.

Semedo, D. 2009. *Credit Scoring: Aplicação da Regressão Logística vs Redes Neurais Artificiais na Avaliação do Risco de Crédito no Mercado Cabo-Verdiano*. Universidade Nova de Lisboa, Lisboa.

Seneviratna, D. 2006. Credit Risk Management In Consumer Finance, *Lankan Banking – A Paradigm Shift*: 59-68.

Sicsú, A. 2010. *Credit Scoring*: Edgar Blucher.

Siddiqi, N. 2006. Scorecards: General Overview, *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*: 5-10. New Jersey: John Wiley & Sons, Inc.

Silva, F. 2011. *A Sustentabilidade Financeira e a Gestão do Risco de Crédito no Microcrédito – O Scoring como uma alternativa com valor para a gestão do risco de crédito: uma experiência para Portugal*. Universidade Técnica de Lisboa, Lisboa.

Siqueira, J. 2000. Risco: da Filosofia à Administração. Disponível em <http://www.hottopos.com/convenit3/josiq.htm> consultado a 21.04.2014: Faculdade de Letras da Universidade do Porto.

Smith, K. 2014. History of Credit Rating Agencies and How They Work. Disponível em <http://www.moneycrashers.com/credit-rating-agencies-history/> consultado a 23-04-2014.

Sousa, A., e Chaia, A. 2000. Política de Crédito: uma análise qualitativa dos processos em empresas. *Caderno de pesquisas em administração*, 7 (3): 13-25.

Stavins, J. 2000. Credit Card Borrowing, Delinquency, and Personal Bankruptcy *New England Economic Review*: 15-30

Stein, R. 2007. Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation *The Journal of Risk Model Validation*, 1 (1): 77-113.

Stiglitz, J., e Weiss, A. 1981. Credit rationing in markets with imperfect information. *American Economic Review*, 71(3): 393-410.

Thomas, L. 2000. A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16 (2): 149-172.

Thomas, L., Edelman, D. e Crook, J. 2002. The History and Philosophy of Credit Scoring, The Practice of Credit Scoring e Statistical Methods for Building Credit Scorecards, *Credit Scoring and Its Applications*: 1-41 e 41-62. Philadelphia: SIAM on Mathematical Modeling and Computation.

VantageScore, S. 2013. VantageScore 3.0 White Paper, *Better predictive ability among sought-after borrowers*: 1-9:
http://www.vantagescore.com/images/resources/VantageScore3-0_WhitePaper.pdf consultado a 10.05.2014.

Wesley, D. 1993. Credit risk management: lessons for success. *The Journal of Commercial Lending* (75): 32-38

West, D. 2000. Neural network credit scoring models. *Computers & Operations Research*, 27: 1131-1152.

White, L. 2010. Markets: The Credit Rating Agencies. *The Journal of Economic Perspectives*, 24 (2): 211-226.

Yobas, M., Crook, J. e Ross, P. 2000. Credit scoring using neural and evolutionary techniques. *IMA Journal of Mathematics Applied in Business and Industry* (11): 111-125.

Zhang, T., e DeVaney, S. 1999. Determinants of consumer's debt repayment pattern. *Consumer Interests Annual* (45): 67-70.

Legislação

Decreto-Lei n.º 133/2009 (2009.Jun.02): 3438-3452: Diário da República, 1.ª série N.º 106.

Memorando de Entendimento sobre as condicionalidades de política económica (2011.Mai.17): 1-35.

Regulamento (CE) n.º 1060/2009: 1-31: Jornal Oficial da União Europeia.

Anexos

Anexo 1 Análise de variáveis comumente incluídas em modelos de notação de risco de particulares (quadro comparativo de vários autores).

Anexo 2 Análise de variáveis comumente incluídas em modelos de notação de risco de particulares utilizando bases de dados relativas a famílias (quadro comparativo de vários autores).

Anexo 3 Lista de agências de *rating* mundiais: detalhe do tipo de notações emitidas e país onde estão registradas.

Anexo 4 Estimação do modelo para as bases de dados imputadas de 1 a 5.

Anexo 3 Lista de agências de *rating* mundiais: detalhe do tipo de notações emitidas e país onde estão registadas.

Lista de Agências de Rating <small>Fonte: ESMA (2014) e DefailRisk.com</small>	Tipos de notações emitidas em 2013						Tipos de notação por Agência
	Empresas	Instituições Financeiras	Seguradoras	Estados	Produtos Financeiros Estruturados	Obrigações	
Agusto & Co. Ltd. (Nigéria, Quênia e Ruanda)							4
Ahbor Rating (Uzbequistão)							2
A.M. Best Company, Inc. (Dubai, Hong Kong, México, Reino Unido e EUA)							2
ARC Ratings S.A. (África do Sul, Brasil, Índia, Malásia, e Portugal)							2
ASSEKURATA Assekuranz Rating-Agentur GmbH (Alemanha)							1
Axesor S.A. (Espanha)							6
BCRA-Credit Rating Agency AD (Bulgária)							3
BRC Investor Services S.A. (Colômbia)							1
Calificadora de Riesgo, PCA (Paraguai e Uruguai)							3
Class y Asociados S.A. Clasificadora de Riesgo (Perú)							2
Capital Intelligence, Ltd. (Chipre e Hong Kong)							3
Capital Standards Rating (CSR) (Kuwait)							5
Caribbean Information & Credit Rating Services Ltd. (CariCRIS) (Caribe)							3
CERVED Group S.p.A. (Itália)							1
Creditreform Rating AG (Alemanha)							3
Credit Rating Agency of Bangladesh, Ltd. (CRAB)(Bangladesh)							4
Credit-Rating: A Ukrainian rating agency (Ucrânia)							3
CRIF S.p.A. (Itália)							1
DagongCredit Rating Srl (China e Itália)							3
Demotech, Inc. (EUA)							1
Dominion Bond Rating Service (DBRS) (Canadá e Reino Unido)							6
Economist Intelligence Unit Ltd (Reino Unido)							1
Egan-Jones Rating Company (EUA)							2
Emerging Credit Rating Ltd (ECRL) (Bangladesh)							4
Euler Hermes Rating GmbH (Alemanha)							1
European Rating Agency a.s. (Eslováquia)							1
Feri EuroRating Services AG (Alemanha)							1
Fitch (Alemanha, China, Colombia, Coreia, Equador, Espanha, França, Itália, Malásia, Paquistão, Perú, Polónia, Reino Unido e EUA)							6
GBB-Rating Gesellschaft für Bonitätsbeurteilung mbH (Alemanha)							1
GlobalRating (Arménia, Azerbaijão, Cazaquistão e Rússia)							3
HR Ratings de Mexico, S.A. de C.V. (México)							6
ICAP Group SA (Grécia)							1
Istanbul International Rating Services, Inc. (Turquia)							2
Japan Credit Rating Agency, Ltd. (JCR) (Coreia, Japão, Paquistão e Turquia)							4
Kobirate Uluslararası Kredi Derecelendirme ve Kurumsal Yönetim Hizmetleri A.Ş. (Turquia)							3
Kroll Bond Rating Agency, Inc. (EUA)							3
Mikuni & Co., Ltd. (Japão)							1
Moody's (Alemanha, Chile, China, Chipre, Coreia, Espanha, França, Índia, Indonésia Perú, Reino Unido, Rússia e EUA)							6
National Information & Credit Evaluation, Inc. (NICE) (Coreia)							2
ONICRA Credit Rating Agency of India, Ltd. (Índia)							1
Pacific Credit Rating (PCR) (Perú)							4
Philippine Rating Services, Corp. (PhilRatings) (Filipinas)							3
P.T. PEFINDO Credit Rating Indonesia (Indonésia)							4
Rapid Ratings International, Inc. (Austrália)							1
Rating and Investment Information, Inc. (R&I) (Japão)							3
Realpoint, LLC (EUA)							3
Saha Kurumsal Yönetim ve Kredi Derecelendirme Hizmetleri A.Ş (Turquia)							2
Scope Credit Rating GmbH (Alemanha)							2
Shanghai Credit Information Services Co., Ltd. (China)							2
Shanghai Far East Credit Rating Co., Ltd. (China)							1
SME Rating Agency of India Limited (SMERA) (Índia)							1
Sociedad Calificadora de Riesgo Centroamericana, S.A. (SCRiesgo) (Costa Rica, Guatemala, Honduras, Nicaragua, Panamá e Salvador)							4
Spread Research SAS (França)							1
Standard & Poor's (Argentina, Bangladesh, Chile, China, França, Índia, Itália, Malásia, Polónia, Reino Unido, Sri Lanka e EUA)							6
TCR Kurumsal Yönetim ve Kredi Derecelendirme Hizmetleri A.S. (Turquia)							4
Thai Rating and Information Services Co., Ltd. (TRIS)(Tailândia)							1
TheStreet.com Ratings, Inc.							2
Veribanc, Inc.(EUA)							1

Total Agências de rating (Mundial)	58					
Frequência Absoluta - Tipos de notação	46	32	21	24	18	12
Frequência Relativa - Tipos de notação	79%	55%	36%	41%	31%	21%
Número médio de tipos notação por agência	3					

Frequência Países	Frequência Absoluta	Frequência Relativa
África do Sul	1	1%
Alemanha	8	7%
Argentina	1	1%
Arménia	1	1%
Austrália	1	1%
Azerbaijão	1	1%
Bangladesh	3	3%
Brasil	1	1%
Bulgária	1	1%
Canadá	1	1%
Caribe	1	1%
Cazaquistão	1	1%
Chile	2	2%
China	6	5%
Chipre	2	2%
Colombia	2	2%
Coreia	4	3%
Costa Rica	1	1%
Dubai	1	1%
Equador	1	1%
Eslováquia	1	1%
Espanha	3	3%
Filipinas	1	1%
França	4	3%
Grécia	1	1%
Guatemala	1	1%
Honduras	1	1%
Hong Kong	2	2%
Índia	5	4%
Indonésia	2	2%
Itália	6	5%
Japão	3	3%
Kuwait	1	1%
Malásia	3	3%
México	2	2%
Nicaragua	1	1%
Nigéria	1	1%
Panamá	1	1%
Paquistão	2	2%
Paraguai	1	1%
Perú	4	3%
Polónia	2	2%
Portugal	1	1%
Quênia	1	1%
Reino Unido	6	5%
Ruanda	1	1%
Rússia	2	2%
Salvador	1	1%
Sri Lanka	1	1%
Tailândia	1	1%
Turquia	5	4%
Ucrânia	1	1%
Uruguai	1	1%
EUA	10	8%
Uzbequistão	1	1%

Anexo 4 Estimação do modelo para as bases de dados imputadas de 1 a 5.

Variável	Regressão probabilística para probabilidade de incumprimento ⁽¹⁾ ⁽²⁾ ⁽³⁾				
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Riqueza Líquida	-1,2E-06 (7,4E-07)	-1,2E-06 (7,7E-07)	-1,5E-06 ** (6,8E-07)	-1,4E-06 ** (6,2E-07)	-1,5E-06 ** (6,5E-07)
Dívidas	5,3E-07 (1,0E-06)	5,4E-07 (1,0E-06)	8,0E-08 (1,8E-06)	5,8E-07 (1,9E-06)	
Rendimento	-6,7E-06 ** (3,1E-06)	-7,6E-06 ** (3,1E-06)	1,8E-06 (4,6E-06)	1,8E-06 (4,6E-06)	2,3E-06 (4,0E-06)
Idade	-0,0084 * (0,0043)	-0,0045 (0,0046)	0,0048 (0,0099)	0,0049 (0,0098)	0,0070 (0,0087)
Escolaridade	-0,0643 (0,1490)	-0,0302 (0,1507)	0,0498 (0,2612)	0,0280 (0,2619)	-0,0306 (0,2542)
Situação profissional	0,1156 (0,1052)	0,1304 (0,1057)			
Deterioração das condições	0,5473 *** (0,1016)	0,5268 *** (0,1018)	0,3953 ** (0,2015)	0,3718 * (0,2014)	0,4732 ** (0,1865)
Número de crianças		0,1289 *** (0,0464)	0,2108 ** (0,0915)	0,2100 ** (0,0921)	0,2320 *** (0,0866)
Recusa de crédito			1,2004 *** (0,1990)	1,1931 *** (0,1995)	1,1021 *** (0,1837)
Duração do emprego atual			-0,0222 ** (0,0105)	-0,0208 ** (0,0106)	-0,0164 * (0,0099)
Propriedade de habitação				-0,1822 (0,2237)	-0,0540 (0,1882)
Constante	-0,8780 *** (0,2270)	-1,1578 *** (0,2546)	-1,6125 *** (0,4933)	-1,5118 *** (0,5020)	-1,822 *** (0,4353)
Número de observações	1.543	1.543	392	392	452
Log	-542,81	-538,89	-134,90	-134,58	-152,38
AIC	1101,62	1095,78	289,80	291,17	324,75
Pseudo R ²	7,89%	8,56%	24,75%	24,93%	23,12%
Wald chi ²	59,02	61,87	68,08	68,15	70,36
Acerto do modelo	87,23%	87,23%	85,46%	85,71%	86,50%

(1) Regressão probabilística para probabilidade de incumprimento das famílias usando a base de dados imputada número 1, do ISFF 2010.

(2) O primeiro valor indica o coeficiente (β) e o valor entre parênteses o desvio-padrão.

(3) * $p < 0,1$ ** $p < 0,05$ *** $p < 0,01$

Variável	Regressão probabilística para probabilidade de incumprimento ⁽¹⁾ ⁽²⁾ ⁽³⁾				
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Riqueza Líquida	-1,2E-06 (7,4E-07)	-1,2E-06 (7,7E-07)	-1,5E-06 (6,8E-07)	** -1,4E-06 (6,2E-07)	** -1,6E-06 (6,4E-07)
Dívidas	4,0E-07 (1,0E-06)	4,3E-07 (1,0E-06)	4,8E-08 (1,8E-06)	5,7E-07 (1,9E-06)	
Rendimento	-7,1E-06 ** (3,1E-06)	-7,9E-06 ** (3,1E-06)	1,7E-06 (4,6E-06)	1,8E-06 (4,6E-06)	2,2E-06 (3,9E-06)
Idade	-0,0085 ** (0,0043)	-0,0046 (0,0047)	0,0047 (0,0099)	0,0048 (0,0098)	0,0070 (0,0087)
Escolaridade	-0,0600 (0,1488)	-0,0262 (0,1504)	0,0470 (0,2631)	0,0240 (0,2642)	-0,0349 (0,2561)
Situação profissional	0,1083 (0,1053)	0,1239 (0,1057)			
Deterioração das condições	0,5462 *** (0,1017)	0,5260 *** (0,1020)	0,3927 * (0,2013)	0,3685 * (0,2011)	0,4696 ** (0,1867)
Número de crianças		0,1300 *** (0,0466)	0,2140 ** (0,0913)	0,2133 ** (0,0919)	0,2353 *** (0,0866)
Recusa de crédito			1,1973 *** (0,1995)	1,1899 *** (0,2000)	1,0990 *** (0,1840)
Duração do emprego atual			-0,0221 ** (0,0105)	-0,0207 ** (0,0106)	-0,0162 (0,0099)
Propriedade de habitação				-0,1857 (0,2257)	-0,0578 (0,1882)
Constante	-0,8592 *** (0,2268)	-1,1453 *** (0,2557)	-1,6055 *** (0,4925)	-1,5030 *** (0,5000)	-1,8140 *** (0,4356)
Número de observações	1.542	1.542	392	392	452
Log	-542,76	-538,78	-134,81	-134,48	-152,21
AIC	1101,52	1095,56	289,63	290,97	324,43
Pseudo R ²	7,88%	8,56%	24,80%	24,99%	23,20%
Wald chi ²	59,44	62,40	68,14	68,23	70,50
Acerto do modelo	87,22%	87,22%	85,46%	85,46%	86,28%

(1) Regressão probabilística para probabilidade de incumprimento das famílias usando a base de dados imputada número 2, do ISFF 2010.

(2) O primeiro valor indica o coeficiente (β) e o valor entre parênteses o desvio-padrão.

(3) * p<0,1 **p<0,05 ***p<0,01

Variável	Regressão probabilística para probabilidade de incumprimento ^{(1) (2) (3)}				
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Riqueza Líquida	-1,2E-06 * (7,4E-07)	-1,2E-06 (7,6E-07)	-1,4E-06 ** (6,4E-07)	-1,3E-06 ** (5,8E-07)	-1,5E-06 ** (6,4E-07)
Dívidas	3,5E-07 (1,1E-06)	3,5E-07 (1,0E-06)	2,0E-08 (1,8E-06)	5,5E-07 (1,9E-06)	
Rendimento	-6,2E-06 ** (3,1E-06)	-7,0E-06 ** (3,1E-06)	1,8E-06 (4,6E-06)	1,8E-06 (4,6E-06)	2,0E-06 (3,9E-06)
Idade	-0,0086 ** (0,0043)	-0,0048 (0,0046)	0,0044 (0,0099)	0,0045 (0,0098)	0,0070 (0,0087)
Escolaridade	-0,0702 (0,1499)	-0,0368 (0,1514)	0,0407 (0,2630)	0,0178 (0,2639)	-0,0359 (0,2559)
Situação profissional	0,1161 (0,1051)	0,1314 (0,1056)			
Deterioração das condições	0,5457 *** (0,1016)	0,5253 *** (0,1019)	0,3950 ** (0,2018)	0,3712 * (0,2016)	0,4709 ** (0,1868)
Número de crianças		0,1268 *** (0,0463)	0,2119 ** (0,0912)	0,2112 ** (0,0918)	0,2299 *** (0,0858)
Recusa de crédito			1,1961 *** (0,1989)	1,1888 *** (0,1994)	1,0986 *** (0,1832)
Duração do emprego atual			-0,0222 ** (0,0105)	-0,0207 * (0,0106)	-0,0164 * (0,0099)
Propriedade de habitação				-0,1895 (0,2251)	-0,0562 (0,1882)
Constante	-0,8694 *** (0,2258)	-1,1461 *** (0,2530)	-1,5959 *** (0,4907)	-1,4923 *** (0,4983)	-1,8125 *** (0,4350)
Número de observações	1.541	1.541	392	392	452
Log	-543,15	-539,36	-539,36	-134,66	-153,00
AIC	1102,30	1096,71	290,01	291,32	325,26
Pseudo R ²	7,79%	8,44%	24,70%	24,89%	0,23
Wald chi ²	58,14	61,14	68,28	68,40	70,64
Acerto do modelo	87,22%	87,22%	85,46%	85,46%	86,50%

(1) Regressão probabilística para probabilidade de incumprimento das famílias usando a base de dados imputada número 3, do ISFF 2010.

(2) O primeiro valor indica o coeficiente (β) e o valor entre parênteses o desvio-padrão.

(3) * $p < 0,1$ ** $p < 0,05$ *** $p < 0,01$

Variável	Regressão probabilística para probabilidade de incumprimento ⁽¹⁾ ⁽²⁾ ⁽³⁾				
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Riqueza Líquida	-1,2E-06 (7,3E-07)	-1,2E-06 (7,5E-07)	-1,4E-06 ** (6,5E-07)	-1,3E-06 ** (5,9E-07)	-1,5E-06 ** (6,3E-07)
Dívidas	3,8E-07 (1,0E-06)	3,9E-07 (1,0E-06)	8,0E-08 (1,8E-06)	6,0E-07 (1,9E-06)	
Rendimento	-6,8E-06 ** (3,1E-06)	-7,7E-06 ** (3,1E-06)	1,8E-06 (4,6E-06)	1,8E-06 (4,6E-06)	2,2E-06 (4,0E-06)
Idade	-0,0085 ** (0,0043)	-0,0047 (0,0046)	0,0046 (0,0099)	0,0047 (0,0098)	0,0070 (0,0087)
Escolaridade	-0,0596 (0,1497)	-0,0253 (0,1513)	0,0422 (0,2625)	0,0196 (0,2633)	-0,0356 (0,2545)
Situação profissional	0,1114 (0,1050)	0,1272 (0,1055)			
Deterioração das condições	0,5430 *** (0,1017)	0,5227 *** (0,1020)	0,3948 * (0,2021)	0,3700 * (0,2020)	0,4705 ** (0,1865)
Número de crianças		0,1291 *** (0,0464)	0,2124 ** (0,0913)	0,2114 ** (0,0919)	0,2324 *** (0,0864)
Recusa de crédito			1,1961 *** (0,1992)	1,1887 *** (0,1996)	1,0981 *** (0,1837)
Duração do emprego atual			-0,0223 ** (0,0105)	-0,0209 ** (0,0106)	-0,0166 * (0,0099)
Propriedade de habitação				-0,1886 (0,2233)	-0,0581 (0,1878)
Constante	-0,8604 *** (0,2267)	-1,1429 *** (0,2548)	-1,6073 *** (0,4934)	-1,5017 *** (0,5021)	-1,8153 *** (0,4362)
Número de observações	1542	1542	392	392	452
Log	-543,15	-539,21	-135,09	-134,74	-152,62
AIC	1102,30	1096,42	290,17	291,49	325,23
Pseudo R ²	7,81%	8,48%	24,65%	24,84%	23,00%
Wald chi ²	58,83	62,12	68,09	68,20	70,34
Acerto do modelo	87,22%	87,22%	85,71%	85,71%	86,28%

(1) Regressão probabilística para probabilidade de incumprimento das famílias usando a base de dados imputada número 4, do ISFF 2010.

(2) O primeiro valor indica o coeficiente (β) e o valor entre parênteses o desvio-padrão.

(3) *p<0,1 **p<0,05 ***p<0,01

Variável	Regressão probabilística para probabilidade de incumprimento ⁽¹⁾ ⁽²⁾ ⁽³⁾				
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Riqueza Líquida	-1,2E-06 (7,2E-07)	-1,2E-06 (7,4E-07)	-1,5E-06 ** (6,9E-07)	-1,4E-06 ** (6,3E-07)	-1,5E-06 ** (6,7E-07)
Dívidas	4,7E-07 (1,0E-06)	5,0E-07 (1,0E-06)	-2,2E-08 (1,8E-06)	4,8E-07 (1,9E-06)	
Rendimento	-7,4E-06 ** (3,1E-06)	-8,4E-06 *** (3,1E-06)	1,9E-06 (4,6E-06)	1,9E-06 (4,6E-06)	1,9E-06 (4,1E-06)
Idade	-0,0084 * (0,0043)	-0,0045 (0,0046)	0,0047 (0,0099)	0,0048 (0,0099)	0,0074 (0,0087)
Escolaridade	-0,0549 (0,1500)	-0,0204 (0,1517)	0,0403 (0,2645)	0,0182 (0,2655)	-0,0287 (0,2557)
Situação profissional	0,1069 (0,1052)	0,1221 (0,1056)			
Deterioração das condições	0,5444 *** (0,1017)	0,5239 *** (0,1020)	0,3930 * (0,2015)	0,3703 * (0,2012)	0,4687 ** (0,1870)
Número de crianças		0,1302 *** (0,0463)	0,2121 ** (0,0915)	0,2116 ** (0,0921)	0,2301 *** (0,0862)
Recusa de crédito			1,2004 *** (0,1994)	1,1935 *** (0,1999)	1,1024 *** (0,1832)
Duração do emprego atual			-0,0222 ** (0,0105)	-0,0208 ** (0,0106)	-0,0166 * (0,0099)
Propriedade de habitação				-0,1777 (0,22619)	-0,0469 (0,1894)
Constante	-0,8584 *** (0,2253)	-1,1427 *** (0,2531)	-1,6045 *** (0,4953)	-1,5059 *** (0,5035)	-1,8289 *** (0,4385)
Número de observações	1541	1541	392	392	452
Log	-542,60	-538,61	-134,90	-134,59	-152,58
AIC	1101,19	1095,21	289,79	291,18	325,15
Pseudo R ²	7,89%	8,57%	24,76%	24,92%	23,02%
Wald chi ²	59,02	62,00	68,01	68,10	70,25
Acerto do modelo	87,22%	87,22%	85,46%	85,46%	86,28%

(1) Regressão probabilística para probabilidade de incumprimento das famílias usando a base de dados imputada número 5, do ISFF 2010.

(2) O primeiro valor indica o coeficiente (β) e o valor entre parênteses o desvio-padrão.

(3) *p<0,1 **p<0,05 ***p<0,01