



NOVA

IMS

Information
Management
School

MEGI

Mestrado em Estatística e Gestão de Informação
Master Program in Statistics and Information Management

Rácio de crédito vencido no crédito à habitação

A sua ligação com os indicadores de política
macroprudencial

Diana Madalena Cunha Martins

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre em Estatística e Gestão de
Informação

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

**RÁCIO DE CRÉDITO VENCIDO NO CRÉDITO À HABITAÇÃO - A SUA
LIGAÇÃO COM OS INDICADORES DE POLÍTICA
MACROPRUDENCIAL**

por

Diana Madalena Cunha Martins

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Estatística e Gestão de Informação, Especialização em Análise e Gestão de Informação

Orientador/Coorientador: Professor Jorge Morais Mendes

Coorientador:

Janeiro 2020

RESUMO

Esta análise tem como objetivo principal descrever, de uma forma sistematizada, a relação do rácio de crédito vencido com alguns dos instrumentos macroprudenciais, limitando esta análise apenas para os particulares, desde 2005 até 2017, verificando até que ponto estes conceitos estão correlacionados. Para que seja possível atingir o objetivo descrito anteriormente foi necessário, primeiramente, utilizar um método de imputação para preencher dados em falta no reporte das instituições. O método utilizado foi o pacote MICE (*Multivariate Imputation By Chained Equations*) que cria múltiplas imputações para os dados através da amostragem de *Gibbs*, gerando valores plausíveis para os dados tendo em conta as relações existentes com o resto das variáveis existentes nos dados. Posteriormente, de forma a proceder à análise dos dados, foi utilizado um método de *clustering* para grandes volumes de dados, o método de *clustering* CLARA (*Clustering Large Applications*) que replica o método PAM (*Partition Around Medoids*) para grandes amostras e foi também utilizada uma regressão que se baseia num modelo de duas partes. Este utiliza um modelo binário para modelar a probabilidade de ser observado um valor positivo comparativamente com a probabilidade ser observado um valor igual a zero e um modelo de regressão apropriado para estimar um resultado positivo.

Ao analisar a relação do rácio de crédito vencido com os instrumentos macroprudenciais descritos anteriormente num primeiro passo, foi possível através do método CLARA agrupar todas as observações com rácio de crédito vencido apenas num *cluster* sem que fosse incluída na análise a variável rácio de crédito vencido. Apenas o período da crise apresentou uma alocação das observações ao *cluster* errado (em cerca de 1,5% das observações). Formados estes dois *clusters* de mutuários, e independentemente do período em observação, é possível verificar que as observações alocadas ao mesmo *cluster* partilham características semelhantes, muito influenciadas pelos instrumentos macroprudenciais, em que num dos *clusters* encontramos observações com critérios de concessão de crédito desejáveis (LTI, DSTI, LTV e TAN abaixo da média) e o oposto no outro.

Num segundo passo, foi possível perceber que as variáveis que tiveram um maior impacto sobre o rácio de crédito vencido foram a Taxa Anual Nominal (TAN) do contrato, o facto de o crédito ser contratado numa instituição classificada como de risco sistémico, o ano em que o contrato foi celebrado e, por fim, a maturidade. Todas estas variáveis demonstram ter um impacto positivo na variável em estudo com exceção da variável que diz respeito ao facto de o crédito ser contratado numa instituição classificada como de risco sistémico. O LTV e o DSTI, que não tiveram um impacto tão forte no rácio de crédito vencido, também demonstraram ter um impacto positivo na variável em estudo.

KEYWORDS

Rácio de crédito vencido; LTV; DSTI; MICE; CLARA; Modelo de duas partes

AGRADECIMENTOS

É com muita felicidade que, no fim da minha dissertação, escrevo estes agradecimentos. É com orgulho que me dirijo às pessoas que me acompanharam ao longo deste ano que se encontram destacadas nos próximos parágrafos.

Primeiramente, gostaria de agradecer e dedicar esse trabalho à minha família, que é, sem dúvida, o pilar mais importante na minha formação educacional e pessoal e, em especial, à minha mãe que por diversas vezes me ajudou nesta dissertação e que me ensinou a nunca desistir e à minha irmã por ser a melhor amiga que alguém pode pedir.

Ao Banco de Portugal por me permitir estudar o tema em questão e pelos dados fornecidos que permitiram estudar o fenómeno em questão.

Ao Professor Doutor Jorge Morais Mendes, meu orientador, por aceitar orientar este trabalho e pelo acompanhamento no desenvolvimento do mesmo.

Um agradecimento especial ao Gonçalo Santos por todo o apoio, companheirismo e incentivo prestado durante a elaboração deste documento e por nunca ter deixado de demonstrar o seu orgulho por mim.

Gostaria também de agradecer ao Daniel Abreu por todas as discussões, comentários e sugestões que tiveram um grande contributo para a construção deste trabalho.

Obrigada a todos os que de uma forma ou de outra contribuíram para a realização deste caminho académico!

As opiniões expressas nesta dissertação são minhas e não refletem as opiniões do Banco de Portugal. Quaisquer erros e omissões são da minha única responsabilidade.

Índice

1. Introdução.....	10
1.1. Problemática	10
1.2. Objetivos	12
1.3. Estrutura da dissertação	12
2. Revisão da Literatura.....	13
2.1. Os determinantes da concessão de crédito	14
2.2. Linhas de defesa das autoridades	15
2.2.1. Medidas macroprudenciais em Portugal	16
2.3. Impactos das medidas macroprudenciais – Experiência noutros países	19
2.3.1. Ásia	21
2.3.2. Dinamarca	22
2.3.3. Noruega	22
2.3.4. República Checa	22
2.3.5. Roménia	23
2.3.6. Suécia	23
2.3.7. Suíça	24
2.3.8. A União Europeia.....	24
2.3.9. Conclusão	25
2.4. Rácio de crédito vencido	27
3. Metodologia	29
3.1. Dados utilizados	30
3.2. Variáveis	31
3.2.1. Estrutura da análise dos dados	31
3.2.2. Tratamento dos dados	32
3.3. Imputação multivariada através de equações em cadeia.....	33
3.3.1. Especificação do modelo de imputação.....	33
3.3.2. Algoritmo do método MICE	35
3.3.3. Análise da qualidade das imputações	36
3.4. Análise exploratória dos dados	39
3.5. Construção dos Clusters.....	40

3.6.	Análise de regressão – Regressão de duas partes	42
3.6.1.	Estimação do modelo	44
3.6.2.	Avaliação dos modelos.....	45
3.6.3.	Qualidade dos modelos.....	45
3.6.4.	Seleção de modelos.....	46
4.	Resultados e discussão.....	48
4.1.	Análise de clusters.....	48
4.1.1.	Antes da Crise.....	48
4.1.2.	Durante a crise	50
4.1.3.	Depois da Crise	52
4.2.	Análise de regressão	54
4.2.1.	Estimação dos modelos.....	54
4.2.2.	Interpretação dos coeficientes de regressão	54
5.	Conclusões.....	60
5.1.	Recomendações futuras.....	61
6.	Bibliografia	63
7.	Anexos	67
7.1.	Anexo I - Glossário de variáveis.....	68
7.2.	Anexo II - Métodos de cálculo	69
7.3.	Anexo III - Imputação multivariada através de equações em cadeia (Package MICE).....	70
7.4.	Anexo IV - Clustering Large Applications (CLARA).....	71
7.5.	Anexo V - Estatísticas descritivas	72
7.6.	Anexo VI - Matriz de correlações	74
7.7.	Anexo VII - Número ótimo de clusters utilizando o método da silhueta média.....	76
7.8.	Anexo VIII - Regressão de duas partes	77

Índice de figuras

Figura 1 - Limites ao Loan-to-Value	17
Figura 2 - Limites ao Debt Service-to-Income	18
Figura 3 - Limites à maturidade	18
Figura 4 - Requisitos de pagamentos regulares.....	19
Figura 5 – Número de países que utilizam limites ao rácio de LTV, DSTI e LTI.....	27
Figura 6 - Taxa de crescimento real do PIB em percentagem	31
Figura 7 - Desenho da Investigação	32
Figura 8 - Gráfico de densidade dos valores imputados para o período de 2005 a 2007	37
Figura 9 - Gráfico de densidade dos valores imputados para o período de 2015 a 2017	37
Figura 10 - Silhueta do melhor resultado obtido com o método CLARA.....	49
Figura 11 - Gráfico de <i>clusters</i> utilizando o melhor resultado obtido com o método CLARA	49
Figura 12 - Silhueta do melhor resultado obtido com o método CLARA.....	51
Figura 13 - Gráfico de <i>clusters</i> utilizando o melhor resultado obtido com o método CLARA	51
Figura 14 - Silhueta do melhor resultado obtido com o método CLARA.....	53
Figura 15 - Gráfico de <i>clusters</i> utilizando o melhor resultado obtido com o método CLARA	53
Figura 16 - Número ótimo de <i>clusters</i> para o período de 2008 a 2014.....	76
Figura 17 - Número ótimo de <i>clusters</i> para o período de 2015 a 2017.....	76

Índice de tabelas

Tabela 1 - Países da Europa que aplicam requisitos de capital, limites ao LTV e limites ao DSTI ou LTI no último trimestre de 2018.....	25
Tabela 2 - Utilização de limites ao rácio de LTV, DSTI e LTI por país	26
Tabela 3 - Comparação das imputações efetuadas com o package <i>MICE</i> para o período 2005 a 2007.....	38
Tabela 4 - Comparação das imputações efetuadas com o package <i>MICE</i> para o período 2015 a 2017.....	38
Tabela 5 – Comparação entre os modelos	46
Tabela 6 - Valores médios dos indicadores para o período de 2005 a 2009	49
Tabela 7 - Valores médios dos indicadores para o período de 2010 a 2014	51
Tabela 8 - Valores médios dos indicadores para o período de 2015 a 2017	53
Tabela 9 – Resultado da estimação do modelo de duas partes	56
Tabela 10 - Estatísticas Descritivas para as variáveis relativas ao período de 2005 a 2007	72
Tabela 11 - Estatísticas Descritivas para as variáveis relativas ao período de 2008 a 2014.....	72
Tabela 12 - Estatísticas Descritivas para as variáveis relativas ao período de 2015 a 2017	73
Tabela 13 - Matriz de correlações para o período de 2005 a 2007	74
Tabela 14 - Matriz de correlações para o período de 2008 a 2014	74
Tabela 15 - Matriz de correlações para o período de 2015 a 2017	75

Lista de abreviações e acrónimos

CCyB	Countercyclical capital buffers
CLARA	Clustering Large Applications
CNB	Czech National Bank
DSTI	Debt-Service-to-Income
DTI	Debt to Income
ESRB	European Systemic Risk Board
FCS	Fully Conditional Specification
FMI	Fundo Monetário Internacional
G20	Grupo dos 20
GLM	Modelo Linear Generalizado
JM	Joint Modeling
LTI	Loan-to-Income
LTV	Loan-to-Value
MAR	Missing at random
MICE	Multivariate Imputation by Chained Equations
MNAR	Missing not at random
NIBUD	National Institute for Family Finance Information
OLS	Ordinary Least Squares
O-SII	Other Sistemically Important Institutions
p.p	pontos percentuais
PAM	Partition Around Medoides
PIB	Produto Interno Bruto
PMM	Predictive Mean Matching
RCV	Rácio de crédito vencido
TAN	Taxa Anual Nominal
UE	União Europeia

1. INTRODUÇÃO

1.1. PROBLEMÁTICA

O mercado imobiliário, em particular o residencial, tem estado associado a importantes crises financeiras internacionais, incluindo a crise global financeira de 2007 a 2009. Foi na sequência desta crise que foi introduzido o acordo de Basileia III que se foca maioritariamente em aumentar a quantidade e qualidade do capital e em introduzir requisitos quantitativos de liquidez para bancos individuais. Em paralelo introduziram-se requisitos macroprudenciais (sobretudo reservas de capital) que têm como objetivo evitar a acumulação de risco sistémico, contribuir de forma sustentável para o crescimento económico, e a longo prazo tornar o sistema financeiro mais resistente aos riscos sistémicos minimizando o impacto na economia. Este estudo dá uma forte importância ao uso de alguns destes indicadores macroprudenciais como o “*loan-to-value*” (LTV), o “*debt service-to-income*” (DSTI) e o “*loan-to-income*” (LTI) como ferramentas macroprudenciais para mitigar o risco de crédito excessivo e o nível excessivo do endividamento do setor privado não financeiro. O rácio de *loan-to-value* é o rácio entre o empréstimo e o valor/preço da casa, o rácio *debt service-to-income* corresponde ao esforço mensal dos mutuários para pagar as suas dívidas e o rácio *loan-to-income* corresponde ao esforço mensal, em anos, dos mutuários para pagar o empréstimo que está a realizar.

As políticas macroprudenciais são um fenómeno recente no mundo de reformas regulatórias na União Europeia (UE), onde todas as decisões tomadas sobre os instrumentos de política macroprudencial são tomadas e implementadas pela autoridade macroprudencial nacional, que em muitos países é o Banco Central.

Sendo o objetivo da política macroprudencial limitar a acumulação de risco sistémico, contribuir de forma sustentável para o crescimento económico, e por fim, a longo prazo tornar o sistema financeiro mais resistente aos riscos minimizando o impacto na economia¹, podemos questionar-nos “Como pode então ser alcançada a estabilidade financeira?” Para responder a esta questão temos de considerar dois contextos distintos. O primeiro, o contexto das instituições financeiras que têm de caminhar para uma gestão sã e prudente em termos de rentabilidade e de solvabilidade, recorrendo-se para tal de métodos de gestão e de controlo de riscos adequados e de uma verificação da solvabilidade das suas contrapartes. Contudo, por vezes a primeira linha de defesa da instituição não é suficiente para evitar que as instituições individualmente assumam riscos excessivos ou adotem critérios de concessão de crédito menos prudentes, pelo que surge a segunda linha de defesa por parte das autoridades de supervisão que têm de implementar medidas para antever ou reduzir as situações de instabilidade financeira, recorrendo a formas mais adequadas para atingir a estabilidade financeira.

¹ Banco de Portugal (2014), “*A política macroprudencial em Portugal: objetivos e instrumentos*”.

É também indispensável que as autoridades atuem o mais cedo possível no ciclo, pois quando a sobrevalorização de preços imobiliários aumenta é difícil ser revertida se esta já se tiver acomodado. Caso esta reversão não aconteça, irá existir uma necessidade de intervenção que provém de falhas de mercado, do facto de existirem instituições individuais que podem fazer transitar externalidades negativas para o sistema, por existir uma grande facilidade de contágio como fonte de propagação dos choques no sistema, e por fim, pelo facto de existirem disparidades de informação. Existem dois casos em que podem existir disparidades de informação, o primeiro ocorre quando existe uma assimetria na informação que antecede o contrato entre um comprador e um vendedor (normalmente é o vendedor o detentor da informação, neste caso estamos perante uma seleção adversa - *adverse selection*). O segundo caso, intitulado de risco moral (*moral hazard*) corresponde ao risco de que uma das partes do contrato não forneça toda a informação que tem disponível, fazendo com que alguma das partes corra um risco desnecessário, por exemplo se o mutuário não fornecer toda a informação sobre a sua capacidade financeira.

Este estudo focar-se-á apenas no mercado imobiliário para o contexto da estabilidade financeira. O mercado imobiliário é um setor bastante importante na economia real, tendo este setor um peso substancial na riqueza das famílias e no balanço das instituições financeiras². Neste sentido, é imprescindível ter políticas preventivas e dinâmicas possíveis de ser ajustadas às circunstâncias de cada momento, uma vez que as bolhas do mercado imobiliário podem tanto ter grandes impactos a nível financeiro, como impactos na estabilidade macroeconómica do país. O único problema é que a implementação de políticas macroprudenciais tem subjacente algo extra desafiante por dois motivos. Primeiro, por requerer uma seleção muito crítica das melhores ferramentas de escolha e monitorização a serem utilizadas, de forma a ser determinado qual o tempo e a intensidade dos ajustamentos que necessitam de ser feitos à política implementada. Em segundo, pelo facto de os mecanismos de transmissão da política macroprudencial ainda não serem conhecidos.

Entre as medidas macroprudenciais existentes³, estão duas que destaco neste estudo: i) a imposição de limites máximos ao rácio de LTV, apenas aplicável a créditos hipotecários; e, ii) a imposição de limites máximos ao rácio de DSTI, aplicável tanto a créditos hipotecários como créditos ao consumo, concedidos pelas instituições financeiras.

A introdução destas medidas, na Europa, ainda se encontra numa fase inicial, sendo por isso reduzido o número de países que aplicou medidas macroprudenciais de forma ativa até aos dias que correm, onde Portugal não é exceção. No entanto, sem a implementação destas medidas a concessão de crédito a mutuários com um perfil de risco mais arriscado poderia ter propensão para crescer fazendo, potencialmente, com que o rácio de crédito vencido aumentasse, aquando se observe uma alteração inesperada no rendimento dos mutuários que leva ao aumento dos riscos. Isto porque se na concessão de crédito não são tidos em conta os fatores relevantes para que os mutuários reúnam as condições necessárias para garantir que conseguem efetivamente amortizar

² European Systemic Risk Board (2019), "*Vulnerabilities in the residential real estate sectors of the EEA countries*"

³ Como por exemplo, a imposição de ponderadores de risco associados a exposições colateralizadas por imóveis (*risk weights*), a imposição de limites no LTV e no DSTI e as exigências de reservas.

o crédito (incluindo capital e juros), o crédito acabará eventualmente por se tornar num crédito vencido, o que prejudica o capital próprio das instituições de crédito e a sua capacidade de financiar o sistema.

1.2. OBJETIVOS

Esta dissertação pretende assim analisar os principais indicadores que contribuem para o rácio de crédito vencido e o nível de impacto dos mesmos no crédito vencido, focando no setor dos particulares, comparando as características das operações de crédito contratadas no período antes da crise (2005-2007), no período da crise (2008-2014) e no período pós-crise (2015-2017). Desta forma irá ser possível perceber se as operações associadas aos créditos vencidos têm elevados valores de LTV, DSTI/LTI e de maturidade, ou seja, perceber se estas medidas estão correlacionadas.

Tendo em vista este objetivo as instituições de crédito foram divididas em dois grupos de instituições. Instituições classificadas pelo Banco de Portugal como de importância sistémica⁴ e o resto das instituições a operar em Portugal. Posteriormente foram analisados quais os indicadores que mais contribuem para o rácio de crédito vencido das instituições de crédito.

1.3. ESTRUTURA DA DISSERTAÇÃO

Incluindo a introdução este trabalho encontra-se estruturado em 5 capítulos.

Um segundo capítulo relacionado com a elaboração teórica relativa às políticas macroprudenciais, a sua definição, contexto e aplicação em diversos países dentro e fora da União Europeia de forma a apoiar a compreensão e análise da dissertação. Um terceiro capítulo, que apresenta a metodologia utilizada na elaboração da dissertação, onde são descritos os dados utilizados e todas as técnicas utilizadas para tratar e analisar os dados. No quarto capítulo, é efetuado todo o estudo empírico e a interpretação de resultados. O quinto capítulo conclui este trabalho, incluindo as limitações do mesmo e as recomendações a serem seguidas em trabalhos futuros.

A escolha do tema justifica-se pela sua importância e por ser um tema relacionado com a atualidade, pois com a recente introdução das medidas macroprudenciais em Portugal, tudo se tornará distinto e está ainda a ser avaliado o seu impacto no sistema financeiro.

⁴ A classificação efetuada pelo Banco de Portugal está disponível em:
<https://www.bportugal.pt/page/reserva-de-o-sii>

2. REVISÃO DA LITERATURA

Durante a crise financeira global de 2007-2008 foi possível identificar na economia diversos tipos de vulnerabilidades, sendo estas: i) o sobredimensionamento das instituições; ii) a determinação inadequada do preço do crédito, como por exemplo a aplicação de valores mais baixos de *spread* (uma componente da taxa de juro que é aplicada pelos bancos nos contratos de crédito e, que pode ser entendido como o nível de risco do mutuário) para certos tipos de operações; iii) o crescimento excessivo do crédito/alavancagem; iv) a existência de uma grande dependência face aos mercados financeiros para obtenção de financiamento; v) o aparecimento de bolhas especulativas, nomeadamente no mercado imobiliário residencial; vi) a existência de um foco muito microprudencial na regulação e na supervisão, em vez de se complementar este foco com uma regulação e supervisão macroprudencial; e, vii) um enfoque demasiado nacional, deixando de parte algumas fontes de risco, como por exemplo as interligações entre os grupos e o risco subjacente ao comportamento coletivo.

Ao serem materializadas tais vulnerabilidades foi possível entender que a ocorrência de períodos de dificuldades no acesso a alguns segmentos de mercado por alguns participantes traz uma consequente penalização em termos de crescimento económico, que suscitou, a nível global, intensa discussão quanto aos fatores que estiveram na sua origem e, conseqüentemente, quanto às formas de corrigir tais vulnerabilidades ou insuficiências.

Segundo o Fundo Monetário Internacional (FMI, 2013), existem três tipos de instrumentos macroprudenciais que ajudam a colmatar diversas vulnerabilidades existentes: i) Reserva contra cíclica (*Countercyclical capital buffers* - CCyB) e provisões; ii) ferramentas sectoriais; e, iii) ferramentas de liquidez. Por um lado, o objetivo das ferramentas sectoriais, que incluem entre outros, os constrangimentos nos rácios de LTV e restrições nos rácios de DSTI, consistem em controlar a acumulação excessiva do risco num setor em particular, por outro lado as reservas contra cíclicas pretendem reter o crescimento excessivo de crédito e aumentar a resiliência do setor bancário em sofrer choques, sendo que são uma ferramenta chave da regulação de Basileia III, onde se definem requisitos significativamente mais altos para absorção de perdas não antecipadas⁵. Adicionalmente as ferramentas de liquidez podem ser eficientes para conter a expansão excessiva de crédito dos bancos.

Segundo Neves (2010), após ter sido identificada a necessidade de modificar a regulação e supervisão do setor financeiro para serem corrigidas debilidades detetadas e fortalecer a estabilidade financeira, e após uma avaliação efetuada pelos mais variados organismos e fóruns internacionais (Grupo dos 20 (G20), FMI, Financial Stability Board, Comité de Basileia, Comissão Europeia, entre outros), foram identificadas inúmeras soluções para que as instituições financeiras

⁵ Basel Committee on Banking Supervision (2010b), “An Assessment of the Long-Term Impact of Stronger Capital and Liquidity Requirements”.

tenham responsabilidades acrescidas e apresentem uma maior solidez protegendo assim os consumidores financeiros.

2.1. OS DETERMINANTES DA CONCESSÃO DE CRÉDITO

Uma vez que as instituições financeiras são responsáveis pela concessão de crédito às famílias é necessário perceber as orientações internas que estas instituições utilizam no processo de concessão de crédito.

Uma vez que o sistema financeiro está muito exposto ao mercado residencial, sobretudo via crédito à habitação, em períodos de queda dos preços de habitação serão as famílias e o sistema financeiro, potencialmente, os mais afetados.

Segundo o Banco Central Europeu (BCE, 2015), os critérios de concessão de crédito de uma instituição refletem um conjunto de termos e condições específicos do negócio de cada banco. Para além do volume de negócios e da taxa de juro existem determinantes como o colateral, a maturidade e o limite do empréstimo que também são determinantes muito importantes no processo de concessão de crédito.

É crucial entender os mecanismos que estão por detrás das decisões que concedem crédito às empresas e às famílias e por isso o BCE realizou um estudo⁶ que incluiu 118 bancos de 16 países da área do euro. Com este estudo foi possível concluir que: i) os critérios de concessão de crédito são cíclicos e tendem a ser menos restritos quando o PIB aumenta e a ser mais restritos numa fase de desaceleração económica; ii) normalmente os níveis mais baixos de taxas de juro de curto prazo tendem a suavizar os critérios de concessão, e, em particular quando se verificam níveis mais baixos de taxas *overnight*; iii) os critérios de concessão de crédito tendem a ser menos restritivos com o aumento da qualidade do colateral do mutuário; iv) para suavizarem os critérios de concessão de crédito as instituições financeiras tendem a reduzir os *spreads* do empréstimo, a reduzir a exigência do colateral ou a aumentar as maturidades.

A preocupação com o risco e com os critérios de concessão de crédito não são algo novo na mentalidade das instituições, estas sempre utilizaram indicadores internos para verificar a sustentabilidade da concessão de crédito a cada mutuário. Esta preocupação é possível de observar quando comparamos o inquérito aos bancos sobre o mercado de crédito realizado pelo Banco de Portugal em janeiro de 2003 e o mesmo inquérito de outubro de 2019. Neste estudo é possível verificar que as instituições financeiras sempre estiveram preocupadas com a qualidade creditícia do mutuário, com a maturidade do contrato e com as garantias exigidas e a sua qualidade (através do rácio entre o valor do empréstimo e o valor da garantia). No entanto, estes indicadores eram flexíveis e poderiam não ser tão restritivos no momento de concessão de crédito, como por exemplo, dando demasiado valor ao colateral dos mutuários ou até estendendo a maturidade do

⁶ <https://www.ecb.europa.eu/pub/pdf/fsr/financalstabilityreview200906en.pdf>

contrato para facilitar os pagamentos mensais dos mutuários, já que não existia nenhuma diretiva do supervisor em contrário.

Segundo o BCE (2015) as facilidades descritas anteriormente em suavizar os critérios de concessão de crédito, tendo em conta a fase do ciclo financeiro, podem levar a um aumento excessivo do risco por parte dos bancos em fases de recuperação económica levando a problemas que se podem materializar em problemas no balanço dos bancos, problemas estes que acabarão por levar as instituições a ser mais restritivas nos seus critérios de concessão tendo um impacto negativo na concessão de crédito, afetando não só uma instituição, mas também poderiam afetar o setor financeiro num todo.

Atualmente a recuperação do crédito é algo considerado como um enorme desafio para as entidades credoras. Ainda assim, o processo de renegociação de crédito é sempre a solução à qual se deve dar prioridade, para assim se chegar a um acordo entre ambas as partes para que seja possível minimizar os custos de incumprimento.

Para tentar evitar problemas no setor financeiro algumas autoridades de diversos países introduziram medidas macroprudenciais para colmatar estes problemas.

2.2. LINHAS DE DEFESA DAS AUTORIDADES

Como referido anteriormente, é necessário considerar duas linhas de defesa quando se pretende responder à questão: “Como pode ser alcançada a estabilidade financeira?”. A primeira linha de defesa diz respeito às instituições financeiras que têm de caminhar para uma gestão sã e prudente em termos de rentabilidade e de solvabilidade, recorrendo-se para tal de métodos de gestão e de controlo de riscos adequados e de uma verificação da solvabilidade das suas contrapartes.

Para assegurar esta gestão sã e prudente, a instituição deve garantir o cumprimento de todos os mecanismos microprudenciais existentes, cujo objetivo não é substituir a gestão competente dos órgãos de administração, nem a gestão de controlo interno dos órgãos de fiscalização e as funções de controlo, nem o papel desempenhado pelos auditores externos das instituições, mas sim assegurar que a instituição seja capaz de salvaguardar a sua própria solvabilidade e solidez financeira e a segurança dos fundos que lhes sejam confiados⁷.

Por outro lado, e, de forma a evitar que exista uma verificação desadequada da solvabilidade dos mutuários e, também, de forma a proteger os consumidores financeiros, o Banco de Portugal emitiu um aviso⁸, direcionado à avaliação de solvabilidade dos mutuários, para que esta análise de solvabilidade seja feita através de certos procedimentos e critérios que devem ser observados no momento da avaliação dos consumidores pelas entidades que concedem crédito.

⁷ <https://www.bportugal.pt/page/micro-objetivos-e-principios?mlid=1231>

⁸ Aviso n.º 4/2017 do Banco de Portugal, que transpõe parcialmente a Diretiva n.º 2014/17/UE, do Parlamento Europeu e do Conselho.

No entanto, esta primeira linha de defesa nem sempre é eficiente, pois não evita que as instituições assumam riscos desnecessários ou que existam choques na economia que nem sempre podem ser previstos. Por estas razões é necessário também que exista uma segunda linha de defesa que é constituída pelas autoridades de supervisão que têm de implementar medidas para antever ou reduzir as situações de instabilidade financeira, recorrendo a formas mais adequadas para atingir a estabilidade financeira. Uma forma que Portugal encontrou para aumentar a resiliência do papel das instituições financeiras na economia do país foi introduzindo uma Medida Macroprudencial⁹, denominada por Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito (de agora em diante denominada por Recomendação do Banco de Portugal).

2.2.1. Medidas macroprudenciais em Portugal

Enquanto autoridade macroprudencial, o Banco de Portugal, considera que, no contexto económico que hoje presenciamos, ainda se verifica um nível elevado de endividamento dos particulares comparativamente com o contexto europeu¹⁰ acompanhado de baixas taxas de juro e de poupança.

A medida macroprudencial visa que as instituições financeiras apenas concedam crédito em situações em que nem as instituições nem as famílias assumam riscos excessivos. Desta forma, o rácio de crédito vencido, uma *proxy* para os riscos que os Bancos aceitam, poderá reduzir no futuro uma vez que as instituições são forçadas a serem mais prudentes nos critérios de concessão de crédito.

A medida anteriormente referida é aplicável aos novos créditos concedidos a partir de 1 de julho de 2018 a consumidores (pessoas singulares que atuam com objetivos alheios à sua atividade comercial ou profissional, em contratos de crédito abrangidos pelo Decreto-Lei n.º 133/2009¹¹ e no Decreto-Lei n.º 74-A/2017¹²). Esta medida aplica-se aos créditos com garantia hipotecária ou equivalente e créditos ao consumo, incluindo recomendações relativamente a: i) limites ao rácio de LTV; ii) limites ao rácio de DSTI; iii) limites à maturidade dos empréstimos; e, iv) requisitos de pagamentos regulares de juros e de capital.

Rácio Loan-to-Value

A introdução de um limite ao LTV corresponde nada mais do que à imposição de um valor máximo ao nível de crédito que pode ser disponibilizado tendo em conta o valor da garantia hipotecária. Desta forma, pretende-se que os mutuários tenham capital próprio para adquirir o bem pretendido.

O LTV é apenas passível de ser calculado para contratos que têm pelo menos um colateral imóvel, o que significa que todos os contratos ao consumo estão excluídos para o cálculo deste indicador.

⁹ Banco de Portugal (2018), “Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito celebrados com consumidores”.

¹⁰ Banco de Portugal (2018), “Relatório de Estabilidade Financeira, junho 2018”.

¹¹ https://www.bportugal.pt/sites/default/files/anexos/legislacoes/267726285_2.doc.pdf

¹² https://www.bportugal.pt/sites/default/files/anexos/legislacoes/350414490_1.doc.pdf

Este indicador corresponde ao rácio entre o montante do empréstimo e o valor do imóvel dado em garantia

$$\frac{\text{Montante de empréstimo(s) garantido(s) pelo mesmo imóvel}}{\text{Min (preço de aquisição; Avaliação do imóvel dado em garantia)}} \quad (1)$$

No valor do denominador é utilizado, por questão de conservadorismo, o menor entre o preço de aquisição e a avaliação do imóvel dado em garantia.¹³

Os limites ao LTV impostos pela Recomendação do Banco de Portugal encontram-se detalhados na figura seguinte

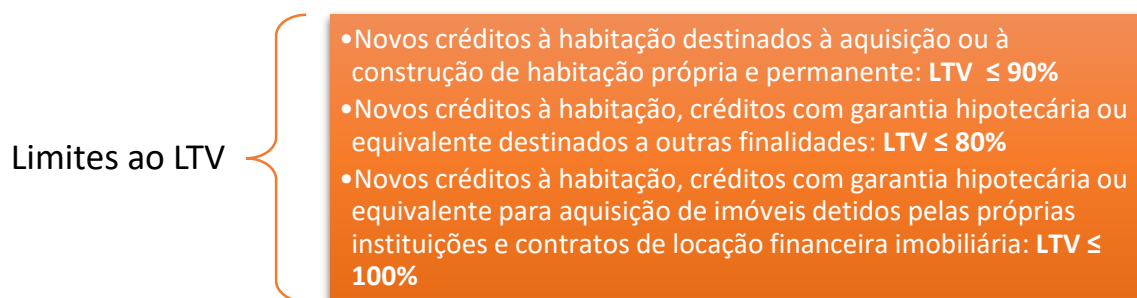


Figura 1 - Limites ao Loan-to-Value

Fonte: Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito celebrados com consumidores, 2018

Em geral, as restrições ao rácio de LTV estão associadas a um efeito de desaceleração relativa ao crescimento dos preços das casas, a transações de imóveis em menores quantidades e à redução no volume de crédito.

Para o FMI (2013) a ferramenta macroprudencial mais utilizada são restrições ao rácio de LTV. Segundo Crowe et al. (2011) em 2011, limites ao LTV foram reportados como utilizados em 36 países num inquérito do FMI que cobriu 60 países.

¹³ Banco de Portugal (2018), “Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito celebrados com consumidores”.

Rácio Debt Service-to-Income

O rácio ao DSTI dá ênfase aos encargos mensais/serviços de dívida associados ao empréstimo. Corresponde ao rácio entre o montante das prestações mensais, calculadas com todos os empréstimos do(s) mutuário(s), e o seu rendimento líquido de impostos e contribuições obrigatórias à Segurança Social.

$$\frac{\text{Todas as prestações de empréstimos detidos pelo(s) mutuário(s)}}{\text{Rendimento mensal líquido do(s) mutuário(s)}} \quad (2)$$

No caso do denominador, sempre que no final do contrato algum dos mutuários envolvidos no empréstimo tenha mais de 70 anos, o rendimento mensal do(s) mutuário(s) deve sofrer uma redução de 20%, durante o período em que algum dos mutuários tenha mais de 70 anos, de forma a que se tenham em conta que no momento da reforma o mutuário terá uma quebra nos seus rendimentos. Também as prestações mensais do novo contrato de crédito devem ser calculadas assumindo que serão constantes ao longo do período em que o contrato esteja em vigor. Por fim, no caso de contratos a taxa de juro variável e mista, deve ser considerado o impacto de um aumento da taxa de juro¹⁴.

Limites ao DSTI

- Os contratos de crédito devem apresentar um DSTI $\leq 50\%$, com as exceções do montante total de créditos concedidos por cada instituição, em cada ano:
 - ate 20% pode ter DSTI até 60%
 - até 5% pode ultrapassar os limites ao DSTI

Figura 2- Limites ao Debt Service-to-Income
Fonte: Recomendação do Banco de Portugal, 2018

Maturidade original dos empréstimos

Os limites à maturidade foram introduzidos de forma a que as instituições de crédito ou os mutuários não tentem reduzir o valor do DSTI dos contratos ao aumentar a maturidade, contornando assim o objetivo da Recomendação. O limite à maturidade pretende também reduzir a maturidade média dos empréstimos em Portugal já que a média que se verifica está muito acima da média histórica de Portugal e da média da UE.

Limites à maturidade

- Nos crédito à habitação, créditos com garantia hipotecária ou equivalente:
 - Maturidade de cada novo contrato ≤ 40 anos
 - Maturidade média do conjunto de novos contratos deve convergir, de forma gradual, para 30 anos, até ao final de 2022
- Nos contratos de crédito ao consumo:
 - Maturidade de cada novo contrato ≤ 10 anos

Figura 3 - Limites à maturidade
Fonte: Recomendação do Banco de Portugal, 2018

¹⁴ Banco de Portugal (2018), *Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito celebrados com consumidores*

Pagamentos regulares

A Recomendação¹⁵ estabelece ainda que os contratos de crédito devem ter pagamentos regulares de capital e juros, uma vez que a crise financeira permitiu concluir que a presença de uma redução significativa no preço das casas durante a recessão e o crescimento significativo dos preços das casas antes da recessão estão possivelmente associados com a grande amplitude da recessão. Sendo assim as mudanças nos preços das casas tendem a ter um papel crítico ao determinar a duração e o custo da recessão.

Requisitos de pagamentos regulares

- Os novos contratos de crédito devem ter pagamentos regulares de capital e juros

Figura 4 - Requisitos de pagamentos regulares
Fonte: Recomendação do Banco de Portugal, 2018

Conclusão

Em suma, ao serem cumpridos todos estes requisitos do Banco de Portugal, espera-se que os créditos sejam concedidos a mutuários com suficiente capacidade financeira para suportar o novo crédito e, caso existam, os créditos anteriores. Assim, o crédito vencido das novas operações poderá ter propensão para decrescer ao longo dos anos uma vez que, em teoria, o crédito é concedido em melhores condições. Sendo assim, parece existir uma correlação entre todos os indicadores de concessão de crédito e o rácio de crédito vencido das instituições.

Segundo o European Systemic Risk Board – ESRB (2019) e o Banco de Portugal (2019) após a implementação da Recomendação existe uma menor percentagem de crédito a ser concedido a mutuários com níveis de DSTI superiores a 60% e LTV superior a 90%, estando a aumentar os créditos a mutuários com DSTI entre 50% e 60% e com LTV entre 80% a 90%.

2.3. IMPACTOS DAS MEDIDAS MACROPRUDENCIAIS – EXPERIÊNCIA NOUTROS PAÍSES

Segundo Neagu *et al.* (2015) os instrumentos de DSTI e LTV mostram ser bastante eficientes em reduzir o rápido crescimento de crédito e assegurar que tanto o devedor como o credor estão capazes de lidar com choques adversos no valor real dos preços, depreciações no valor da moeda ou picos de taxas de juro. Sendo assim os instrumentos de LTV e DSTI parecem ser os instrumentos macroprudenciais mais eficazes a deter o rápido crescimento de crédito.

Segundo Bernanke (2010), os instrumentos macroprudenciais são mais eficazes do que os instrumentos de política monetária, uma vez que a política monetária não tem um alcance tão grande como a política macroprudencial ao ponto de ajudar a estabilizar o mercado imobiliário.

¹⁵ Banco de Portugal (2018), *Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito celebrados com consumidores*

Segundo Neagu *et al.* (2015) existem algumas mensagens para melhorar a eficácia ao utilizar os instrumentos de DSTI e LTV. A primeira prende-se com o facto de ser necessário fortalecer a cooperação entre a autoridade macroprudencial doméstica e as outras autoridades estrangeiras para que se preserve a eficácia destes instrumentos ao lidarem com riscos. A segunda tem a ver com o facto de ser mais prudente monitorizar o nível total do endividamento dos particulares, em vez de apenas se focar no endividamento doméstico dos particulares.

Com o propósito de estudar o impacto dos limites sobre o LTV no setor imobiliário Almeida *et al.* (2005) encontraram evidência de que os preços à habitação e os novos empréstimos são mais sensíveis a choques no rendimento em países onde os limites ao LTV são superiores, já que são os países com um limite ao LTV mais elevado que apresentam uma capacidade da dívida positivamente correlacionada com variações no ciclo económico, e um mercado de financiamento imobiliário mais desenvolvido.

Crowe *et al.* (2011) estendem a sua análise a 36 economias que verificaram bolhas no setor imobiliário e consideraram que 24 tomaram medidas de política monetária, fiscal ou macroprudencial a partir de setembro de 2000, demonstrando que medidas macroprudenciais mais direcionadas, como é o caso do LTV e do DSTI, estão associadas a um nível de sucesso superior comparativamente com outras políticas.

Kuttner e Shim (2013) analisam 9 tipos de políticas que remontam até 1980 e incluem um total de 60 países. Estes concluem que existe evidência de que reduções incrementais nos limites ao DSTI afetem mais o crescimento do crédito hipotecário do que limites ao LTV, diminuindo o crescimento do crédito em 4 a 7 pontos percentuais (p.p) nos quatro trimestres seguintes. As reduções aos limites ao DSTI são mais efetivas do que os limites ao LTV pois durante explosões nos preços das casas, o aumento do seu preço faz com que o valor máximo que pode ser emprestado aumente também, fazendo com que os limites ao LTV sejam mais voláteis. No entanto estes resultados são bastante sensíveis ao modelo econométrico escolhido.

Arregui *et al.* (2013) utilizam variáveis macroprudenciais para extrapolar o impacto médio da aplicação de restrições mais exigentes ao nível das diferentes políticas macroprudenciais, chegando à conclusão de que as políticas mais restritivas, como o LTV o DSTI e os ponderadores de risco, levam não só à redução do crescimento do rácio entre o crédito e o Produto Interno Bruto (PIB) mas também à redução dos preços de mercado de habitação.

Krznar e Morsink (2014) reproduziram a metodologia de Arregui *et al.* (2013) para 25 países que introduziram medidas quantitativas para reduzir o crescimento do crédito hipotecário e o preço das casas no período de 2000 a 2012. Estes encontraram evidência de que as políticas mais restritivas ao nível dos LTV's, DSTI's e ponderadores de risco estão associadas a um crescimento inferior no crédito hipotecário entre 0,5 e 0,75 pontos percentuais durante o período em que são aplicadas.

Segundo Neagu *et al.* (2015), existe uma forte correlação negativa entre o nível de endividamento e a capacidade de pagar a dívida, ou seja, quando aumenta o endividamento, a capacidade de pagar a dívida diminui. Os devedores com rendimentos baixos e DSTI altos tendem a ter um rácio de

crédito vencido alto, independentemente do propósito do empréstimo (ao consumo ou para crédito hipotecário). O nível de LTV está também relacionado com a capacidade do mutuário pagar a dívida, ou seja, os devedores com LTV elevados tendem também a ter um rácio de crédito vencido alto. Existe ainda evidência de que a aplicação destas medidas por parte dos bancos sem a intervenção das autoridades apresenta resultados piores.

Montalvo *et. al* (2017) afirmam que as condições financeiras do crédito hipotecário são importantes, podendo o banco tentar compensar o maior risco dos empréstimos com LTV mais elevados cobrando *spreads* superiores. Concluem também que as instituições com maior apetite por risco, indicado pelos seus standards na origem do empréstimo, apresentam maiores níveis de LTV, demonstrando que a maior parte dos créditos hipotecários foram aprovados independentemente do perfil de risco dos mutuários. A hipótese de Montalvo *et. al* (2017) é que se a sobrevalorização via preço de avaliação foi usada para melhorar o *score* do crédito, e serão as instituições financeiras com maior apetite, que utilizam esta prática, por risco que podem vir a ter uma intervenção do Banco de Espanha, ou com uma injeção de capital público.

2.3.1. Ásia

A Ásia utiliza medidas macroprudenciais já há largos anos sem lhes chamar especificamente medidas macroprudenciais, mas nem todos os países asiáticos obtêm o mesmo resultado.

Segundo Se (2013), em 21 países da Ásia, os limites de LTV estão positivamente relacionados com a apreciação dos preços entre 2000 e 2007. Verificou-se que reduções nos limites do LTV levaram a uma diminuição na apreciação dos preços dos imóveis. Por exemplo, na Coreia do Sul, onde os limites ao LTV foram introduzidos em 2002 seguidos por limites ao DSTI em 2005, a taxa em cadeia dos preços da habitação desceu de 3,4% para 0,3% imediatamente após à introdução do LTV permanecendo baixos até Abril de 2003. Subsequentes reduções ao LTV levaram a diminuições significativas na apreciação dos preços dos imóveis.

No caso de Singapura (Monetary Authority of Singapore, 2011), implementaram-se políticas macroprudenciais centradas no mercado de imóveis para tornar possível a existência de um mercado sustentável onde os preços se possam movimentar em linha com os objetivos económicos fundamentais de forma a encorajar uma maior prudência financeira nos compradores e para manter os *standards* da oferta.

Segundo Se (2013), Hong-Kong, onde estes limites já vêm a ser aplicados há 20 anos, têm vindo a confiar bastante nos rácios de LTV, que os têm ajudado a prevenir bolhas de mercado imobiliário reduzindo e mitigando os custos sociais desnecessários de se ter de instalar uma segunda política. Estas políticas não são algo novo, antes da crise de 1997 já eram aplicadas políticas macroprudenciais para proteger o setor bancário dos riscos das bolhas de mercado. Ao longo dos últimos anos, Hong-Kong tem vindo a ganhar experiência nos picos de preços dos imóveis utilizando as lições que aprenderam em 1997 como referência de forma a reconhecer e detetar pressões de mercado, monitorizando os riscos e vulnerabilidades. Hong-Kong desenvolveu também modelos

para avaliar os preços dos imóveis, estando estes a ser utilizados para estimar num longo período de tempo o equilíbrio dos preços dos imóveis baseando-se nos fundamentos da procura e oferta.

2.3.2. Dinamarca

A Dinamarca tem vindo a implementar diversas políticas macroprudenciais desde 2012 que assentam essencialmente em 3 medidas: i) medidas baseadas no capital (*Capital-based measures*); ii) medidas de crescimento de crédito (*Credit growth measures*); e, iii) medidas baseadas nos mutuários (*Borrower-based measures*).

As medidas baseadas no capital foram implementadas entre 2017 e 2018 com o objetivo de aumentar a resiliência das instituições. Para tal a Dinamarca implementou o CCyB que foi ativado com um rácio de 0,5% com início em março de 2019, aumentando mais 1% a partir de setembro de 2019. Com o objetivo de consolidar os *standards* de crédito, que se verificavam ser muito baixos, de forma a limitar o crescimento excessivo de crédito, foram implementadas em 2010 e 2014, onde foi implementado um limite nos empréstimos dos bancos onde o crescimento dos empréstimos a cada segmento de empréstimos (considere-se segmento de empréstimos os seguintes segmentos particulares, propriedade de arrendamento, agricultura e outras empresas não financeiras) de 15% por ano. Por fim as medidas baseadas nos mutuários foram aplicadas entre 2012 e 2017, onde os limites ao Debt- to-Income (DTI) e ao LTV permitiram solidificar os *standards* de crédito e amortecer a capacidade dos mutuários a curto tempo de forma a aumentar a sua resiliência, tanto do setor bancário, como dos mutuários.

2.3.3. Noruega

Segundo Jin et. al (2014), na Noruega, como em muitos outros países, o mercado imobiliário tem sido caracterizado pela elevada alavancagem existente e por isso os rácios de LTV têm aumentado ao longo da última década. Este rácio que se situa agora nos 85% tem como objetivo aumentar a sustentabilidade dos mutuários do mercado imobiliário, estando a ser ponderado medidas alternativas para o complementar. Uma alternativa para complementar o LTV, que implementaram em 2017, foi o LTI, já que este tem a vantagem de medir mais diretamente a sustentabilidade dos mutuários. Jin et. al (2014) constataram também que apesar do rácio de propriedade no mercado imobiliário não se ter modificado, este rácio aumentou entre as pessoas mais jovens, sugerindo que os rácios de LTV mais elevados induziram à procura de habitação por parte de mutuários com rendimentos menores.

2.3.4. República Checa

Segundo Mark Zeman (2018)¹⁶, o Banco Nacional da República Checa (CNB) utiliza vários instrumentos de política macroprudencial, sendo que relativamente aos limites ao LTV impõe que: i) o rácio de LTV não deve ultrapassar os 90%; e ii) os empréstimos com LTV de 80%-90% podem representar não mais de 15% de todos os créditos à habitação concedidos pelo banco. Quanto aos

¹⁶ Disponível em: <https://www.cnb.cz/en/cnb-news/press-releases/CNB-introduces-new-mortgage-limits/>

limites ao DSTI, o CNB impõe que o limite máximo do rácio do DSTI seja 45%. O CNB utiliza ainda, desde 1 de outubro de 2018, o DTI que não deve ser superior a 9. Os bancos podem ainda ultrapassar os limites de DSTI e DTI se esse excesso não exceder 5% dos seus empréstimos. A partir de julho de 2019, o CNB, irá aumentar também o CCyB para 1,5%, que foi introduzido para acautelar os períodos de maior dificuldade. Uma vez que existe uma certa vulnerabilidade dos bancos a uma deterioração na situação económica.

2.3.5. Roménia

O Banco Nacional da Roménia está a considerar redesenhar as suas ferramentas atuais de política macroprudencial, expandindo-as de tal forma que cubram todos os empréstimos das famílias. Estão a ponderar recalibrar os limites existentes ao DSTI, especificamente para: i) incluir um máximo ao limite DSTI que se aplique ao limite stressado do DSTI (após serem impostos choques às taxas de juro e ao rendimento, ao invés de confiarem nos limites internos dos bancos); ii) expandir a aplicabilidade da ferramenta para cobrir todos os empréstimos às famílias (créditos hipotecários, e créditos ao consumo); e, iii) recalibrar os choques para o cálculo do DSTI stressado e introduzir choques similares aplicáveis ao crédito ao consumo.

Segundo o FMI (2018), ao ser analisada a central de registos de crédito da Roménia a probabilidade de incumprimento de um mutuário é muito sensível a qualquer modificação no DSTI e nos rácios de DSTI perto dos 50%, em particular nos créditos à habitação. É por isso importante nunca definir limites acima deste sensível limite.

Para além disso, impor um limite de 50% no limite ao DSTI conduz a um menor rácio de crédito vencido, enquanto o impacto de uma restrição no volume de empréstimos seria limitado. A análise feita pelo FMI (2018) sugere que para os três anos futuros os rácios de crédito vencido serão inferiores entre 0,3 a 0,4 p.p.

Os resultados da regressão logística, utilizada pelo FMI, permite verificar que: i) um aumento de 10 pontos percentuais no DSTI aumenta a probabilidade de incumprimento em 6p.p para créditos à habitação e 3p.p para crédito ao consumo; e ii) o aumento da maturidade residual em um ano leva a um aumento de 20% na probabilidade de incumprimento no crédito ao consumo e um aumento de 4% nos créditos hipotecários.

2.3.6. Suécia

A autoridade macroprudencial da Suécia introduziu um limite de 85% ao rácio de LTV em 2010, no entanto, segundo o FMI (2017), é necessário que a autoridade macroprudencial introduza um valor máximo para o rácio de DTI de forma a aumentar a resiliência das famílias e travar os valores elevados de endividamento. Ao introduzir este limite as famílias passam a ter uma maior margem para absorver o impacto de choques nas taxas de juro já que não se podem endividar aos níveis que hoje se verificam.

Em 2015, o Banco central da Suécia, Riksbank, efetuou um estudo sobre os efeitos do DTI. Segundo o *Copenhagen Economics* (2017), este estudo permitiu concluir que, a curto tempo, se verificará

uma redução os preços da habitação em 4% após a introdução do limite do DTI, mas apenas 0,3% a longo tempo, no entanto, existe incerteza sobre quais os preços à habitação que serão afetados. Encontram também evidência de que o rácio de DTI pouco decresça a curto tempo já que apenas novos empréstimos são afetados pelo limite ao mesmo, e que assim que este limite seja totalmente implementado é previsto que o rácio de DTI diminua 11 pontos percentuais.

2.3.7. Suíça

Segundo o Basel Committee on Banking Supervision (2010a), o CCyB, protege o setor bancário de efeitos prejudiciais do ciclo financeiro. Segundo Basten e Koch (2015), a Suíça, que foi o primeiro país a ativar este *buffer*, em fevereiro de 2013, afirma existir evidência de que o sistema financeiro ficou mais resiliente. Afirmam também que este *buffer* afetou a composição da concessão de crédito, uma vez que o CCyB impõe requisitos de capital mais altos a alguns segmentos (por exemplo habitação). Foram então os bancos que aplicaram restrições de capital que aumentaram mais a sua concessão de crédito comparativamente com os seus concorrentes.

Depois do CCyB ter sido ativado, os bancos especializados, que operam com um modelo de negócios muito intensivo em créditos aumentaram também as suas taxas para um grau superior em termos relativos. Uma vez que o CCyB se aplica a novos créditos, bem como *stocks* de todos os créditos em balanço, os suíços perceberam que o CCyB tende a transferir novos empréstimos hipotecários de bancos relativamente pior capitalizados para bancos relativamente melhores capitalizados, e de bancos relativamente mais expostos a crédito hipotecário para bancos menos expostos a crédito hipotecário. Desta forma alcançam o segundo objetivo da política macroprudencial, ou seja, alocar novo crédito a bancos mais resilientes.

Por fim, os suíços concluíram que: i) os bancos geralmente pedem compensação extra para garantir riscos mais elevados, não parecendo que estas ponderações de risco amplifiquem os efeitos do CCyB nas taxas de crédito e na criação de crédito; ii) a ativação do CCyB contribuiu para a resiliência geral dos bancos, para o aumento das taxas de crédito, especialmente para bancos mais expostos; e, iii) foram os bancos mais resilientes que aumentaram os seus empréstimos. Estes desenvolvimentos suportam os dois objetivos da política macroprudencial, especialmente o segundo objetivo.

2.3.8. A União Europeia

Segundo o ESRB (2019) é possível verificar na tabela seguinte que, no último trimestre de 2018, de 31 países da Europa eram 19 os países que aplicavam limites ao rácio do LTV, 14 os que utilizavam limites ao rácio de DSTI ou ao rácio de LTI e todos utilizavam requisitos de capital como medidas macroprudenciais. No entanto, à data, ainda existiam 11 países da Europa que não aplicavam limites ao LTV ou limites ao DSTI. Esta análise apenas demonstra a importância que as medidas macroprudenciais apresentam nos dias que correm.

Tabela 1 - Países da Europa que aplicam requisitos de capital, limites ao LTV e limites ao DSTI ou LTI no último trimestre de 2018

	Requisitos de Capital	LTV	DSTI/LTI
Áustria	x		
Bélgica	x		
Bulgária	x		
Croácia	x		
Chipre	x	x	x
República Checa	x	x	x
Dinamarca	x	x	
Estónia	x	x	x
Finlândia	x	x	
França	x		
Alemanha	x		
Grécia	x		
Hungria	x	x	x
Islândia	x	x	
Irlanda	x	x	x
Itália	x		
Letônia	x	x	
Lituânia	x	x	x
Luxemburgo	x		
Malta	x		
Holanda	x	x	
Noruega	x	x	x
Polónia	x	x	x
Portugal	x	x	x
Roménia	x	x	x
Eslováquia	x	x	x
Eslovénia	x	x	x
Espanha	x		
Suécia	x	x	
Reino Unido	x		x
Liechtenstein	x	x	

2.3.9. Conclusão

Foram muitos os países que implementaram instrumentos macroprudenciais desde a crise, sendo que já eram alguns os que aplicavam instrumentos macroprudenciais antes. Os países que utilizam o CCyB encontram-se detalhados na 1ª linha da tabela 2. Os países que utilizam limites ao LTV encontram-se na mesma tabela, na linha 2, e por fim, os países que utilizam limites ao DSTI (incluindo limites ao LTI) encontram-se na 3ª linha da tabela 1.

Tabela 2 - Utilização de limites ao rácio de LTV, DSTI e LTI por país

	América	Europa	Oceânia	Ásia	África	Total
CCyB	Peru (2017)	Dinamarca (2019), Eslováquia (2017), França (2016), Islândia (2017), Lituânia (2015), Noruega (2017), Reino Unido (2017), República Checa (2017), Suécia (2015)	Austrália (2016)	Hong Kong (2017)	Mauritânia (2018), Nicarágua (2018)	13
Limites ao rácio de LTV	Bahamas (2017), Brasil (2013), Canadá (2008), Chile (2012), Colômbia (2000), Guatemala (1995)	Áustria (1994), Bielorrússia (2018), Bulgária (2004), Chipre (2016), Eslováquia (2018), Estónia (2015), Finlândia (2010), Holanda (2012), Hungria (2015), Irlanda (2015), Islândia (2017), Letónia (2007), Líbano (2015), Lituânia (2011), Noruega (2010), Polónia (2013), Portugal (2018), República Checa (2017), Roménia (2004), Suécia (2010)	Nova Zelândia (2013)	Arábia Saudita (2014), Butão (2014), China (2001), Coreia (2002), Emirados Árabes Unidos (2013), Hong Kong (1991), Índia (2010), Indonésia (2016), Israel (2012), Kosovo (2016), Líbano (2008), Malásia (2010), Mongólia (2008), Omã (2008), Paquistão (2003), Roménia (2004), Sérvia (2011), Singapura (2010), Sri Lanka (2015), Tailândia (2003), Turquia (2016)	Namíbia (2016), Tanzânia (2015), Tunísia (2016), Uganda (2016)	52
Limites ao rácio de DSTI (incluindo limites ao LTI)	Bahamas (2017), Canadá (2012), Colômbia (2008), Chile (2000)	Bielorrússia (2018), Chipre (2013), Eslováquia (2017), Estónia (2015), Grécia (2005), Holanda (2013), Hungria (2015), Letónia (2013), Lituânia (2011), Noruega (2017), Polónia (2010), Portugal (2018), Reino Unido (2014), República Checa (2018), Roménia (2004)		Arábia Saudita (2018), Bangladesh (2004), Butão (2014), China (2004), Coreia (2005), Egipto (2016), Emirados Árabes Unidos (2011), Hong Kong (1997), Israel (2013), Kosovo (2016), Koweit (1996), Malásia (2011), Mongólia (2008), Omã (2012), Paquistão (2009), Singapura (2013), Tailândia (2017)	Maurícia (2014)	37

Fonte: Base de dados do FMI

Nota: Dentro dos parenteses está o ano em que o país introduziu as medidas.

Segundo os dados recolhidos da base de dados do FMI, a utilização do CCyB, limites ao rácio do LTV e limites ao rácio de DSTI é frequentemente utilizada separadamente, mas por vezes é utilizada em conjunto. Num levantamento que envolveu os 62 países que constam na tabela anterior, relativo à utilização destas 3 ferramentas, é possível de verificar, na figura abaixo, que existem: i) 13 que utilizam requisitos de capital (21% dos países em estudo); ii) 52 países que aplicam limites ao rácio de LTV (84% dos países em estudo); iii) 37 países que aplicam limites ao rácio de DSTI (60% dos países em estudo); iv) 30 países que aplicam limites ao rácio de LTV e ao rácio de DSTI (48% dos países em estudo); v) 33 países utilizam mais do que uma das ferramentas (53% dos países em estudo); e, vii) 4 países utilizam as três ferramentas (6% dos países em estudo).

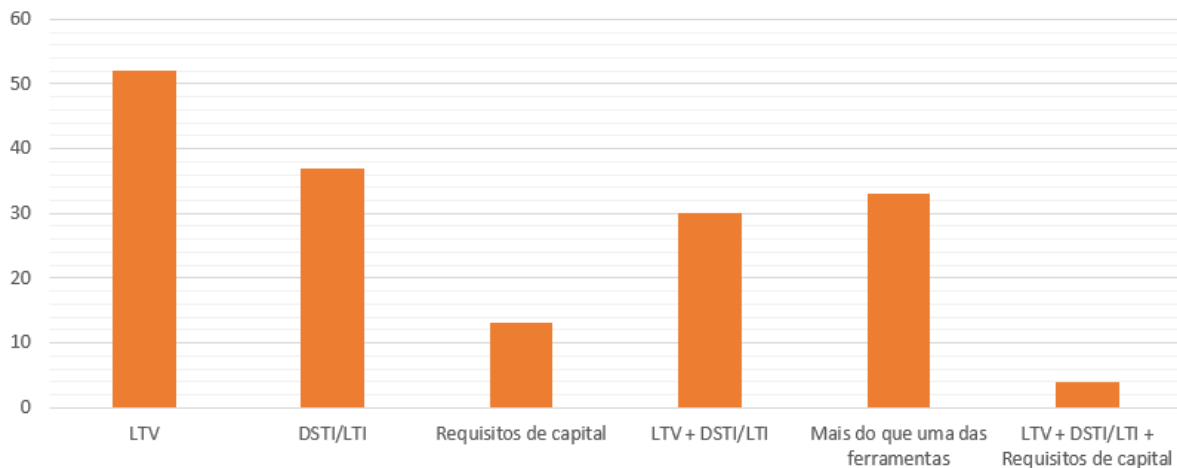


Figura 5 – Número de países que utilizam limites ao rácio de LTV, DSTI e LTI

2.4. RÁCIO DE CRÉDITO VENCIDO

Com a situação que se tem vindo a presenciar nos últimos anos o incumprimento/rácio de crédito vencido registado no crédito concedido a clientes demonstrou que reforçar a sua monitorização e o controlo é imprescindível. Segundo Pereira *et al.* (2015), o volume de crédito vencido das instituições demonstrou ter uma correlação positiva com as dificuldades financeiras sentidas pelas famílias e por empresas, tendo, por isso, passado a ser considerada em muitos estudos, como uma *proxy* da solidez dos bancos.

O aumento do rácio de crédito vencido nas instituições pode indiciar falhas na política de crédito dos bancos, contribuindo, caso seja generalizada, para o aumento da probabilidade de ocorrência de crises financeiras. Com o intuito de atenuar a instabilidade, muitos países transformaram a banca numa das indústrias mais regulamentadas. Mesmo assim, em muitas economias verifica-se não ser suficiente para prevenir os choques sistémicos mais recentes.

Segundo o ESRB (2019) a acumulação excessiva de crédito vencido no sistema financeiro pode ter um impacto na resiliência do setor financeiro. O crédito vencido pode também estar relacionado com uma taxa de financiamento mais elevada e uma menor oferta de crédito, o que pode levar a um sentimento negativo face aos bancos com níveis mais elevados de crédito vencido, diminuindo assim a capacidade dos bancos de aceder ao mercado de capitais. As políticas macroprudenciais podem ter um papel bastante importante em prevenir um problema sistémico de crédito vencido e/ou a aumentar a resiliência do setor financeiro caso este problema se materialize.

No caso de Portugal, contrariamente a alguns países da área do euro, como por exemplo Espanha e Irlanda, não se observou uma sobrevalorização nos preços de habitação no período que procedeu a crise financeira de 2008¹⁷. No entanto, segundo o Banco de Portugal (2017) os preços da habitação

¹⁷RODRIGUES, Raulo & LORENÇO, Rita (2015). "House prices: bubbles, exuberance or something else? Evidence from euro area countries."

caíram substancialmente entre 2010 e 2013 o que levou a um aumento do nível de incumprimento no pós crise, no entanto este manteve-se reduzido, por um lado pelo facto do crédito à habitação ter sido concedido a mutuários com menor risco e, por outro lado, devido ao regime de taxas de juro mais usual ser o de taxas variáveis.

De acordo com o guia de indicadores de solidez financeira do FMI (2004), o crédito torna-se vencido ou em incumprimento se: i) o pagamento de juros ou amortização de capital estão vencidos há pelo menos 90 dias; ii) o pagamento de juros a 90 ou mais dias foi capitalizado, refinanciado ou atrasado por meio de acordo entre partes; iii) os pagamentos estão atrasados há menos de 90 dias mas existem outras razões que induzem a que os mesmos possam não ocorrer na sua totalidade, designadamente a situação de falência do devedor.

De notar que o critério do vencimento do crédito a 90 dias para identificação do crédito vencido é frequente, mas não universal.

Sendo assim, neste estudo para calcular o Rácio de Crédito Vencido irá ser utilizada a seguinte fórmula para os créditos que se encontrarem em incumprimento

$$\frac{\textit{Montante em dívida do crédito em incumprimento}}{\textit{Total de Crédito Concedido}} \quad (3)$$

3. METODOLOGIA

Uma vez que o tema principal deste estudo é a relação entre os indicadores macroprudenciais e o rácio de crédito vencido em três períodos distintos, antes, durante e após a crise, é importante conhecer quais os indicadores que mais contribuíram para este rácio nos diferentes períodos de tempo. Assim, será possível perceber se as instituições financeiras modificaram os critérios de concessão de crédito aquando um evento de crise.

Como os dados utilizados dependem de variáveis que não são obrigatórias no reporte das instituições financeiras ao Banco de Portugal, existem células de algumas variáveis que não se encontram preenchidas na sua totalidade. Destaca-se, em particular, a informação reportada relativa ao rendimento bruto do mutuário na data de celebração do contrato e o valor da primeira prestação do contrato cujo reporte é mais fraco para a generalidade dos bancos. Estas são variáveis essenciais para o cálculo dos indicadores macroprudenciais, que são a essência deste estudo, e por isso foi utilizada uma técnica de imputação multivariada de forma a preencher estas células omissas com valores realistas.

O método utilizado para este estudo, e descrito na secção 3.2, é o método de imputação multivariada através de equações em cadeia (de agora em diante denominado por MICE). Após a utilização desta técnica foi utilizada uma metodologia usualmente aplicada aos estudos de *data mining*, a metodologia de análise de *clusters*. A análise de *clusters*, é apenas um nome vago que na sua essência engloba uma grande variedade de metodologias de estatística multivariada cujo objetivo passa por organizar a informação existente sobre um grupo de entidades em grupos o mais homogêneos possível com base em semelhanças/diferenças entre as unidades em análise. Existem diversos critérios para atingir estes objetivos sendo que os mais conhecidos se baseiam na soma dos quadrados dos erros. Uma alternativa a estes métodos mais conhecidos, e utilizada neste estudo, é a utilização do método dos k-medoides que procura os k objetos mais “representativos” na base de dados, denominados como medoides, que minimizam as dissemelhanças médias dos objetos da base de dados ao medoide mais próximo. Numa primeira iteração estes medoides são as observações mais centrais da base de dados e a cada iteração existente estes vão-se modificando até serem identificados os medoides que efetivamente minimizam a soma das dissemelhanças dos objetos associados a esse mesmo medoide. Um *cluster* é, neste caso, nada mais do que um conjunto de observações associadas ao mesmo medoide.

O método utilizado neste estudo, e detalhado na secção 3.4, é o método CLARA (*Clustering Large Applications*), que se baseia em medoides e que apenas é uma extensão do método dos k-medoides para grandes conjuntos de dados (formados por mais do que algumas centenas de observações), de forma a reduzir tempos de processamento e problemas de memória já que não utiliza toda a base de dados mas utiliza uma amostra da mesma.

Por fim, com o objetivo de descrever a relação entre a variável de interesse (variável dependente), o rácio de crédito vencido, e restantes variáveis (variáveis explicativas) é utilizada uma análise de regressão de duas partes, com o objetivo de acomodar o massivo número de zeros que existem na

variável dependente. Este método utiliza um modelo binário (*probit* ou *logit*) para modelar a probabilidade de ser observado um valor positivo comparativamente com a probabilidade ser observado um valor igual a zero para modelar a primeira parte e um modelo de regressão para estimar um resultado positivo no caso de este se verificar - método dos Mínimos Quadrados Ordinários (OLS) ou um Modelo Linear Generalizado (GLM) – para modelar a segunda parte do modelo. Foi selecionado o melhor método com base nas técnicas descritas na secção 3.6.2 e 3.6.3, que permitem avaliar a qualidade dos modelos.

3.1. DADOS UTILIZADOS

Este estudo foi realizado com os dados recolhidos da instrução n.º19/2016¹⁸, fornecidos pelo Banco de Portugal, devidamente anonimizados para que não fosse possível a identificação da instituição financeira, dos mutuários e dos contratos de crédito.

Os dados apenas contêm informação sobre a classificação das instituições em dois grupos: i) instituições consideradas pelo Banco de Portugal como de importância sistémica¹⁹; e, ii) as restantes instituições financeiras.

A análise efetuada refere-se ao conjunto de contratos celebrados entre 2005 e 2017 relativos a empréstimos destinados à aquisição, construção ou realização de obras em habitação própria permanente, secundária ou para arrendamento, bem como para a aquisição de terrenos para construção de habitação própria (crédito à habitação) e ainda aos contratos garantidos por uma hipoteca que incidem, total ou parcialmente, sobre um imóvel que garante simultaneamente um contrato de crédito à habitação celebrado com a mesma instituição de crédito (crédito conexo).

O período em análise, 2005 a 2017, encontra-se dividido em três períodos de tempo de forma a analisar diferentes comportamentos dos mutuários e das instituições financeiras antes, durante e após um choque na economia. Replicando a técnica de Blanco *et al.* (2018), que utiliza o PIB real para determinar o período da crise Espanhola, é possível observar na figura seguinte que se regista uma variação real do PIB negativa entre os períodos de 2008 a 2014, demonstrando um evento de crise na economia Portuguesa. Sendo assim, os períodos em análise serão: i) operações celebradas no período antes da crise (2005-2007); ii) operações celebradas no período da crise (2008-2014); e, iii) operações celebradas no período pós-crise (2015-2017), período este que não permitirá avaliar os impactos da implementação da Recomendação.

¹⁸ Instrução n.º 19/2016, disponível em: <https://www.bportugal.pt/instrucao/192016-0>

¹⁹ <https://www.bportugal.pt/page/reserva-de-o-sii>

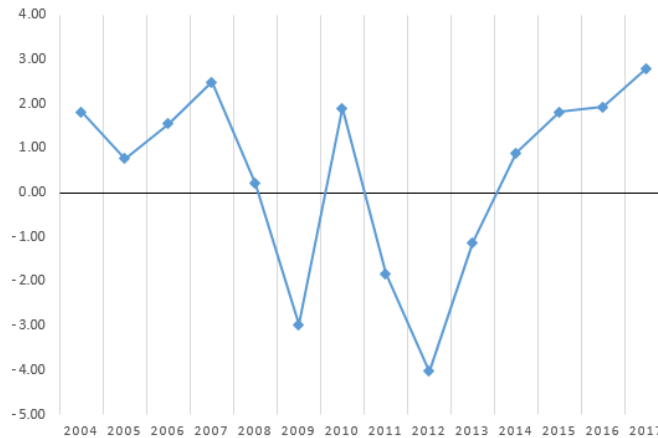


Figura 6 - Taxa de crescimento real do PIB em percentagem
Fonte: PORDATA

3.2. VARIÁVEIS

As variáveis utilizadas neste estudo podem ser repartidas em cinco dimensões: i) grupo de instituições em que foi solicitado o crédito; ii) características do imóvel dado como colateral; iii) condições dos mutuários no momento da contratação; iv) condições do contrato no momento da contratação; e, por fim v) características do contrato em 2017.

Para a primeira dimensão é utilizada uma variável que permite analisar se o crédito foi solicitado numa instituição considerada de risco sistémico (O-SII). As variáveis da segunda dimensão estão relacionadas com o imóvel dado como colateral (PreçoAqHipotecado, ValorPrimeiraAval, LTV). Na terceira dimensão, as variáveis contêm informação sobre as condições do mutuário aquando a solicitação do crédito (DSTI, LTio e RendAnualBrutoDataIni). Na dimensão das condições do contrato no momento da contratação estão incluídas as variáveis que contêm informação sobre o contrato de crédito realizado entre o mutuário e a instituição de crédito (AnoCelebração, Maturidade, MntInicial, MntDivida, ValorPrimeiraPrest e TAN). Por fim, na dimensão características do contrato em 2017 estão incluídas as variáveis relacionadas com as características do contrato à data de reporte do mesmo (Incumprimento_2017, DifAnocelAnoref e RCV)

Para um melhor entendimento de cada uma das variáveis utilizadas neste estudo existe um glossário de variáveis no Anexo I deste documento.

3.2.1. Estrutura da análise dos dados

A análise da informação foi efetuada em três etapas: imputação multivariada, análise exploratória dos dados e por fim a análise de dados, que inclui a análise de *clusters* e a análise de regressão.

Uma vez que o reporte de algumas variáveis essenciais para o cálculo destes indicadores não são de reporte obrigatório na instrução nº19/2016²⁰, foi utilizado um método de imputação multivariada

²⁰ Instrução disponível em: <https://www.bportugal.pt/instrucao/192016>

utilizando o MICE package em R para que os dados omissos das variáveis críticas para o cálculo dos diversos indicadores deste estudo fossem preenchidas com valores plausíveis.

A análise exploratória dos dados foi utilizada para ganhar conhecimento sobre os dados de forma a entender as relações existentes entre as diversas variáveis e para ganhar conhecimento sobre os dados de forma a criar novas variáveis relevantes para o estudo.

A análise de *clusters* foi utilizada para agrupar os mutuários com base em características comuns das variáveis em análise de forma a conhecer melhor os perfis de concessão de crédito, tentando verificar se os contratos com crédito vencido têm características semelhantes. Para este passo foram utilizadas todas as variáveis descritas no Anexo I à exceção da variável RCV.

Por fim, a análise de regressão permite quantificar o impacto de cada uma das variáveis explicativas na variável dependente, o rácio de crédito vencido, de forma a perceber quais as variáveis que mais contribuem para este rácio. Neste passo não foram utilizadas quatro das variáveis descritas no Anexo I (AnoCelebração, MntDivida, Incumprimento_2017 e LTlo) para evitar problemas de multicolinearidade.

Na figura seguinte está ilustrado o desenho da investigação



Figura 7 - Desenho da Investigação

3.2.2. Tratamento dos dados

Aos dados recolhidos foram aplicados alguns filtros de forma a excluir valores que não fizessem sentido ser analisados ou que fossem valores extremos. Foram removidos os contratos em que se detetou:

- Indexante diferente da Euribor, Spread = 0 ou Taxa anual nominal inferior a 0,1% - Contratos entre as instituições financeiras e os seus colaboradores.
- Contratos identificados como conexos garantidos por imóveis sem qualquer associação a contratos de crédito à habitação.
- Preço de aquisição ou valor da primeira avaliação do imóvel inferior a 5.000 euros ou superior a 5.000.000 euros.
- Rendimento anual do(s) mutuário(s) inferior a 2.500 euros ou superior a 1.000.000 euros.
- Valores da primeira prestação inferiores ao primeiro percentil ou superiores ao percentil 99,5.
- Variável LTV superior a 300%.

Após a recolha e tratamento, os dados foram utilizados para calcular os indicadores macroprudenciais em análise e o rácio de crédito vencido contrato a contrato. Estes indicadores foram calculados de acordo com as fórmulas apresentadas no Anexo II.

3.3. IMPUTAÇÃO MULTIVARIADA ATRAVÉS DE EQUAÇÕES EM CADEIA

Antes de ser possível prosseguir com a análise exploratória dos dados através do R e STATA, foi necessário utilizar um método de imputação para preencher células omissas no reporte para apenas dois dos três períodos em análise. Para o período da crise não foi necessário recorrer a uma imputação de valores uma vez que não existia nenhuma célula omissa.

O método de imputação escolhido foi o método MICE, este método utiliza valores plausíveis que são retirados de uma distribuição especialmente criada para cada célula sem valor. Neste método são criadas múltiplas imputações (com reposição de valores) para os dados em falta. Este método é baseado na especificação de condicionalidade total, onde cada variável é imputada por um modelo separado. O algoritmo pode também imputar variáveis contínuas ou categóricas em conjunto.

Para algumas resoluções rápidas, uma substituição pela média pode ser útil, no entanto ao utilizar uma resolução deste tipo acabamos por enviesar os dados uma vez que esta substituição mantém a média igual, mas diminui a variância o que não é de todo desejável.

A imputação multivariada (Rubin, 1987) é o método de escolha para problemas com dados incompletos complexos. Os dados omissos que ocorrem em mais do que uma variável apresentam um desafio especial e neste sentido surgiram duas abordagens gerais para imputação de dados multivariados a modelagem conjunta (Joint Modeling - JM) e a especificação totalmente condicional (Fully Conditional Specification - FCS), também conhecida como imputação multivariada por cadeia de equações (MICE).

A imputação multivariada é uma estrutura que utiliza várias versões imputadas dos dados, substituindo os valores em falta por valores de dados plausíveis. Estes valores plausíveis são extraídos de uma distribuição modelada especialmente para cada célula omissa. Para que os valores sejam o mais próximo possível da realidade o modelo de imputação deve acautelar o padrão dos dados omissos, preservar as relações nos dados e ao mesmo tempo preservar a incerteza sobre essas mesmas relações.

3.3.1. Especificação do modelo de imputação

A especificação do modelo de imputação é o passo mais desafiante na imputação multivariada, já que, segundo Buuren e Groothuis-Oudshoorn (2011) existem sete escolhas essenciais que necessitam de ser efetuadas por parte do utilizador.

1. Devemos decidir se a suposição de falta ao acaso (Missing at random - MAR) é plausível.

Existem dois tipos de dados omissos:

- MAR valores omissos de forma aleatória. Este é o cenário desejável em um caso de dados omissos.

- MNAR (Missing not at random) valores omissos de forma não aleatória. A falta de dados não aleatórios é uma questão mais séria e, neste caso, pode ser aconselhável verificar o processo de recolha de dados e tentar entender por que as informações estão faltando.

O package MICE pode lidar com ambos os MAR e MNAR, no entanto a imputação multivariada sob MNAR requer um modelo diferente. No âmbito deste trabalho assumiu-se sempre a hipótese de que os dados são MAR.

2. É necessário escolher a forma do modelo de imputação, este tem de ser especificado para cada coluna incompleta nos dados.

É necessário escolher o modelo de imputação que se quer utilizar para cada variável que necessita de imputação, uma vez que para cada tipo de variável existem diferentes métodos de imputação. Já que todas as variáveis que apresentam valores omissos neste estudo são variáveis contínuas a opção tomada foi utilizar a função *predictive mean matching* (PMM). Esta função implementa a média preditiva, um método de imputação semi-paramétrico. As suas principais vantagens prendem-se com o facto de as imputações serem restritas aos valores observados, do método conseguir manter o padrão da distribuição da variável e, com o facto de poderem preservar relações não-lineares, mesmo que a parte estrutural do modelo de imputação esteja errada. Isto porque a função PMM retira valores que já existem na distribuição da variável para preencher as células omissas.

3. A terceira escolha diz respeito ao conjunto de variáveis a incluir como preditores na imputação modelo.

O conselho geral é incluir o maior número possível de variáveis relevantes, incluindo as interações (Buuren e Groothuis-Oudshoorn, 2011). Ao usar todas as informações disponíveis a imputação produzida terá um enviesamento menor. Este princípio implica que se deve escolher o maior número possível de preditores, incluindo tantos preditores quanto os possíveis de forma a tornar a suposição do MAR mais plausível. Para fins de imputação, é conveniente seleccionar um subconjunto de dados que contém não mais que 15 a 25 variáveis. Buuren *et al.* (1999) fornecem a seguinte estratégia para seleccionar variáveis a partir de uma grande base de dados:

- Incluir o máximo de variáveis que irão ser utilizadas no modelo que irá ser aplicado aos dados após a imputação. Não o fazer pode influenciar a análise, especialmente se o modelo de dados completos contiver fortes relações preditivas.
- Além disso, deverão ser incluídas variáveis que expliquem uma quantidade considerável de variação. Tais preditores ajudam a reduzir a incerteza das imputações.

É também necessário remover as variáveis que tenham muitos valores ausentes dentro do subgrupo de casos incompletos. Newman (2014) sugere que a percentagem de casos observados incompletos idealmente não ultrapasse os 10%. Neste estudo não existe nenhuma variável com mais de 10% das células omissas.

4. É também necessário decidir se deveríamos imputar variáveis que dependem de outras variáveis incompletas.

Muitos conjuntos de dados contêm variáveis transformadas e com dependências de outras variáveis presentes nos dados. Pode ser útil incorporar as variáveis transformadas no algoritmo de imputação múltipla. Se, por acaso, tanto a variável original como a variável transformada necessitarem de imputações, não existe uma certeza de que as transformações se manterão entre as duas variáveis após a imputação. Existe, no entanto, um mecanismo, denominado de imputação passiva, que permite lidar com tais situações e manter a consistência entre diferentes transformações dos dados, no entanto para o processo de imputação não foi utilizada nenhuma variável transformada ou que dependa de outras.

5. A quinta escolha diz respeito à ordem em que as variáveis devem ser imputadas.

Existem inúmeras estratégias e cada uma com seus respectivos prós e contras. Por *default* o algoritmo do MICE imputa as colunas incompletas nos dados da esquerda para a direita. Teoricamente a ordem de imputação não é relevante, no entanto, pode existir uma ordem mais eficiente que outra. Independentemente da ordem, cada coluna irá ser visitada mais do que uma vez na mesma iteração, o que pode ser útil para assegurar uma sincronização correta entre as variáveis.

6. A sexta escolha diz respeito à configuração das imputações iniciais e ao número de iterações.

Um valor maior de iterações é mais provável de ser eficiente em grandes amostras de dados mas pode ter uma performance mais fraca em amostras pequenas. A literatura (Morris *et al.* 2014) recomenda 10 iterações.

7. A sétima escolha é escolher o número correto para a variável m , o número de imputações distintas que serão efetuadas.

Configurar um m muito baixo pode resultar num grande erro de simulação, especialmente se a percentagem de informação omissa for elevada. Também a ausência de muitos dados pode ser um problema. Schafer (1997) afirma que 3 a 5 imputações são suficientes e Fichmand e Cummings (2003) que 10 são mais do que suficientes. Para este estudo escolheu-se utilizar 7 imputações.

3.3.2. Algoritmo do método MICE

Para cada variável V_i com valores omissos, o processo de equações em cadeia pode ser explicado em 4 passos:

1. Efetua-se uma imputação de valores que resultam da escolha do modelo de imputação escolhido pelo utilizador. Neste caso, uma vez que foi escolhido o PMM, será implementada a média preditiva para cada valor omissa na base de dados. Para cada observação omissa o método PMM aleatoriamente escolhe uma observação que se assemelhe à omissa mas que esteja preenchida para preencher este valor omissa. Esta semelhança depende da semelhança de todas as outras variáveis que estão associadas à observação omissa.

2. As imputações efetuadas no passo 1 para a primeira variável da lista voltam a omissas;
3. É feita uma regressão com todas as variáveis do modelo, em que a variável V_i é a variável dependente e todas as outras variáveis do modelo são utilizadas como variáveis explicativas. Estas regressões operam sujeitas às mesmas suposições de uma análise de regressões, por exemplo, linear, logística ou de *Poisson* fora do contexto de dados omissos.
4. Os valores omissos para a variável V_i são substituídos por imputações do modelo de regressão. Posteriormente quando a variável V_i é utilizada como variável independente na regressão de outras variáveis, serão utilizados tanto os valores observados como os valores imputados.
5. Os passos 2 a 4 são repetidos para cada variável com valores omissos. Uma iteração/imputação encontra-se concluída quando todas as variáveis com valores omissos são sujeitas aos passos 2 a 4.

O *software* utilizado para a análise de *clusters* foi o R (3.5.1) e os *inputs* necessários para utilizar o método MICE estão descritos no Anexo III.

3.3.3. Análise da qualidade das imputações

Do método escolhido resultaram sete imputações distintas. Para analisar a qualidade das imputação e de forma a escolher a melhor imputação recorreu-se primeiramente a gráficos de densidade dos valores imputados comparativamente com os valores reportados, para verificar se de facto os valores são "valores plausíveis", e posteriormente a uma análise das estatísticas descritivas dos dados originais versus as estatísticas descritivas das sete imputações.

Ao analisar os gráficos de densidade dos valores imputados comparativamente com os valores reportados no Anexo III é possível verificar que a densidade dos dados imputados (a magenta) são muito semelhantes à densidade dos dados originais, sendo difícil verificar diferenças significativas na distribuição dos valores imputados, com exceção da variável *RendAnulBrutoDataIni* para o período depois da crise. Adicionalmente recorreu-se a uma comparação das estatísticas descritivas das sete imputações efetuadas com as estatísticas descritivas dos dados originais de forma a escolher a imputação que menos impacto tivesse nestas estatísticas. As estatísticas utilizadas foram o mínimo, a mediana, a média e por fim, o máximo.

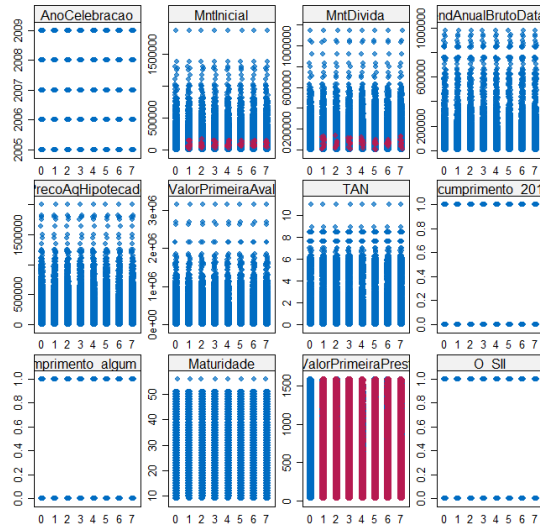


Figura 8 - Gráfico de densidade dos valores imputados para o período de 2005 a 2007

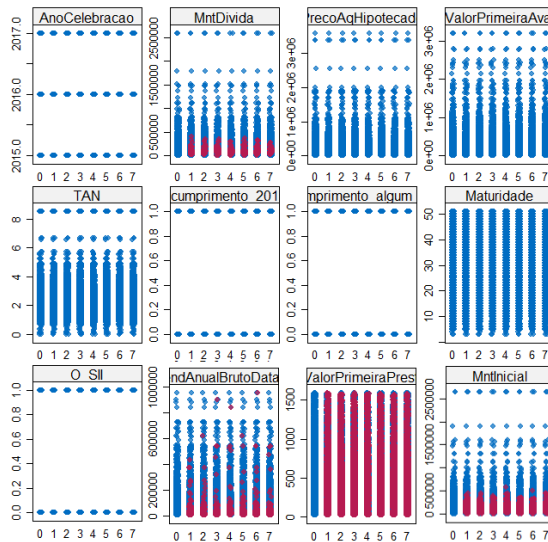


Figura 9 - Gráfico de densidade dos valores imputados para o período de 2015 a 2017

Analisando os valores da tabela 3, é possível observar que a imputação que mais se assemelha aos dados originais é a imputação número 1 e por isso esta foi a imputação utilizada para completar os dados do período antes da crise. Relativamente ao período depois da crise é possível verificar, na tabela 4, que a imputação que mais se assemelha aos dados originais e a que foi utilizada para completar os dados deste período é a imputação 3.

Tabela 3 - Comparação das imputações efetuadas com o package MICE para o período 2005 a 2007

Variável	Estatística	Imputação 1	Imputação 2	Imputação 3	Imputação 4	Imputação 5	Imputação 6	Imputação 7	Original
MntInicial	Mínimo	1 114	1 114	1 114	1 114	1 114	1 114	1 114	1 114
	Mediana	85 000	85 000	85 000	85 000	85 000	85 000	85 000	85 000
	Média	91 849	91 851	91 852	91 851	91 852	91 850	91 850	91 851
	Máximo	1 167 187	1 167 187	1 116 187	1 167 187	1 167 187	1 167 187	1 167 187	1 167 187
MntDivida	Mínimo	76,2	76,2	76,2	76,2	76,2	76,2	76,2	76,2
	Mediana	62 843,5	62 843,6	62 843,4	62 843,6	62843,6	62 843,3	62 843,3	62 843,5
	Média	66 426,01	66 427,9	66 429,0	66 428,2	66429,0	66 427,7	66 426,5	66 428,1
	Máximo	827 831,7	827 831,7	827831,7	827 831,7	827 831,7	827 831,7	827 831,7	827 831,7
ValorPrimeira Prest	Mínimo	28,0	28,0	28,0	28,0	28,0	28,0	28,0	28,0
	Mediana	393,5	393,6	393,7	393,7	393,7	393,6	393,7	394,6
	Média	431,5	432,6	432,7	432,6	432,7	432,5	432,5	430
	Máximo	1 592,4	1 592,4	1 592,4	1 592,4	1 592,4	1 592,4	1 592,4	1 592,4

Tabela 4 - Comparação das imputações efetuadas com o package MICE para o período 2015 a 2017

Variável	Estatística	Imputação 1	Imputação 2	Imputação 3	Imputação 4	Imputação 5	Imputação 6	Imputação 7	Original
MntInicial	Mínimo	570	570	570	570	570	570	570	570
	Mediana	92 975	92 800	92 800	92 975	92 800	92 823	92 800	92 565
	Média	108 392	108 398	108 376	108 459	108 330	108 390	108 377	108 268
	Máximo	2 649 000	2 649 000	2 649 000	2 649 000	2 649 000	2 649 000	2 649 000	2 649 000
MntDivida	Mínimo	0	0	0	0	0	0	0	0
	Mediana	89 003	89 006	89 004	89 002	89 004	89 016	89 004	89 026
	Média	103 146	103 148	103 155	103 122	103 149	103 156	103 147	103 162
	Máximo	2 588 150	2 588 150	2 588 150	2 588 150	2 588 150	2 588 150	2 588 150	2 588 150
ValorPrimeira Prest	Mínimo	14,03	14,03	14,03	14,03	14,03	14,03	14,03	14,03
	Mediana	343,57	343,8	344,31	343,34	344,13	343,77	343,82	342,62
	Média	391,01	391,08	391,35	391,7	391,51	391,23	392,02	387,44
	Máximo	1593,68	1593,68	1593,68	1593,68	1593,68	1593,68	1593,68	1593,68
RendAnualBruto DataIni	Min.	5 006	5 006	5 006	5 006	5 006	5 006	5 006	5006
	Mediana	23 193	23 201	23 209	23 209	23 199	23 201	23 205	23249
	Média	35 211	35 252	35 264	35 267	35 228	35 205	35 197	35222
	Max	955 693	955 693	955 693	955 693	955 693	955 693	955 693	955693

3.4. ANÁLISE EXPLORATÓRIA DOS DADOS

Após a imputação foi possível calcular as variáveis LTVop, LTVoa, DSTI, LTlo e RCV, em STATA, com base nas fórmulas descritas no Anexo II.

Durante a fase de exploração dos dados, devido a um aumento no conhecimento sobre o conjunto de dados em análise, considerou-se importante a criação de uma variável do tipo numérica que corresponde à diferença de anos entre a data de celebração do contrato e a data de referência dos do reporte dos dados (DifAnocelAnoref).

Foi, então, efetuada uma análise descritiva para cada um dos períodos em análise (período antes da crise, período da crise e período depois da crise). Para tal, em primeiro lugar foram calculadas as estatísticas descritivas (Anexo V) mais importantes para cada variável (valor mínimo, médio, mediana, máximo e desvio padrão) de forma a resumir, sumarizar e explorar o comportamento dos dados. De seguida foi efetuada uma análise de matriz de correlações (Anexo VI) para identificar variáveis altamente correlacionadas. Considerou-se que o nível de corte seria uma correlação $\geq |0,8|$.

Através da análise do Anexo V é possível de observar que as variáveis Maturidade, DifAnocelAnoref e LTV apresentam uma distribuição assimétrica negativa, onde a média é inferior à mediana, para os três períodos de análise. A variável O-SII apresenta também uma distribuição simétrica negativa no período antes e durante a crise. No período pós-crise as variáveis Incumprimento_2017 e O-SII apresentam uma distribuição simétrica, onde a média é igual à mediana. Todas as outras variáveis apresentam uma distribuição assimétrica positiva, em que a média é superior à mediana.

É, também, possível de observar que das 17 variáveis apenas 6 apresentam um desvio-padrão alto, independentemente do período em análise. Dado que estas variáveis caracterizarem os indivíduos que solicitam o crédito, faz sentido que a população que pede crédito não seja homogénea. Todas as outras variáveis, que apresentam um desvio-padrão mais baixo, referem-se às características de concessão de crédito das instituições que aparentam ser, num todo, bastante homogéneas.

Com base na matriz de correlações (Anexo VI) é possível de perceber que existem 5 pares de variáveis que apresentam um grau de correlação positiva elevada²¹, $|p| > 0,8$, nos três períodos em análise: i) MntInicial e MntDivida; ii) MntInicial e PreçoAqHipotecado; iii) AnoCelebração e DifAnoCelebAnoRef; iv) Incumprimento_2017 e RCV; e, por fim, vi) DSTI LTlo.

²¹ Quando uma das variáveis aumenta a outra aumenta também.

Todas estas correlações são naturalmente esperadas. O montante em dívida dependerá sempre do montante emprestado que por sua vez dependerá do preço do imóvel que se pretende adquirir. A terceira correlação também é esperada, já que a variável RCV apenas é preenchida se a variável Incumprimento_2017 tomar o valor 1. Por fim, uma vez que o DSTI e LTI são indicadores que pretendem avaliar a mesma realidade e por isso é naturalmente esperado que estes dois indicadores estejam relacionados.

A estes 5 pares de variáveis, no período pós-crise, juntam-se os pares PreçoAqHipotecado e ValorAvaliação; e o par PrecoAqHipotecado e MntDivida. Estes dois novos pares de variáveis permitem perceber que neste período o preço do imóvel passou a estar mais correlacionado com a avaliação do imóvel e com o montante em dívida do que nos períodos anteriores.

3.5. CONSTRUÇÃO DOS CLUSTERS

A fase de construção dos *clusters* tem como principal objetivo agrupar os contratos que possuam características semelhantes. Desta forma será possível analisar as características dentro de cada *cluster* formado.

Uma vez que a base de dados é constituída por 74 665 observações para o período antes da crise, 65 639 observações para o período da crise e 27 727 observações para o período pós crise, foi escolhido, de entre os métodos de *clustering* existentes na bibliografia, como já mencionado, o método de *clustering* para grandes aplicações, devido ao facto de ser um método que pode ser utilizado em grandes conjuntos de dados.

O método CLARA, introduzido por Kaufman e Rousseeuw (1986) é uma extensão do método de *k*-medoides (Partitioning around medoides - PAM) utilizado para grandes conjuntos de dados. O método CLARA não armazena uma matriz de dissemelhanças, utiliza apenas uma amostra da base de dados onde aplica um algoritmo PAM, em que cada observação da base de dados é atribuída ao medoide que esteja mais perto de forma a gerar um número ótimo de medoides para a amostra. A qualidade dos medoides formados é medida através de uma média de dissemelhanças entre os objetos da base de dados completa e o medoide do *cluster* ao qual foi atribuído. O processo é repetido durante um número de vezes previamente definido de forma a minimizar o enviesamento dos dados.

O método CLARA segue os seguintes passos²²:

- Do dataset original são criados, aleatoriamente, vários subconjuntos de dados com um tamanho fixo (*samplesize*).
- É aplicado o algoritmo PAM a cada subconjunto e são escolhidos os *k* objetos representativos (medoides) iniciais. Cada observação da base de dados é atribuída ao medoide que esteja mais perto. O objetivo do algoritmo PAM é encontrar *k clusters* baseados em observações da base de dados, que sejam representativos dos mesmos

²² <http://www.sthda.com/english/wiki/print.php?id=236>

(*medoides*) de forma a distribuir todas as observações pelos k clusters formados. Os objetos representativos para cada *cluster* são escolhidos, inicialmente, aleatoriamente e em cada iteração é calculado o custo do medoide escolhido ao ser substituído por uma nova observação da base de dados de forma a minimizar este custo.

- O custo da substituição do medoide escolhido por uma nova observação é calculado com base na média das dissemelhanças das observações ao medoide que está mais perto. Esta medida é utilizada para que seja possível perceber a boa qualidade do processo de *clustering* das observações.
- É retido o subconjunto cuja média das dissemelhanças é menor.

Formados os clusters, a sua qualidade pode ser avaliada com base no coeficiente da silhueta. Este coeficiente mede o quão bem uma observação foi atribuída ao respetivo *cluster* e estima a média das distâncias entre clusters. O gráfico da silhueta traduz uma medida de quão perto os objetos de um *cluster* estão dos objetos dos clusters vizinhos.

Para cada observação i , a silhueta S_i é calculada da seguinte forma²³

- Para cada observação i , é calculada uma dissemelhança média a_i entre i e os outros pontos do cluster ao qual esta observação pertence.
- Para todos os *clusters* C , aos quais i não pertence, é calculada uma dissemelhança média $d(i, C)$ de i a todas as observações de C . O menor valor de $d(i, C)$ é definido como $b_i = \min_C d(i, C)$. O valor de b_i pode ser visto como a dissemelhanças entre i e o seu cluster vizinho.
- Finalmente, o valor da silhueta da observação i é definido pela seguinte fórmula $S_i = (b_i - a_i) / \max(a_i, b_i)$.

Observações com um valor de S_i elevado (perto de 1) estão bastante bem atribuídas ao seu cluster enquanto observações com um S_i perto de zero podem não estar tão bem atribuídas já que existe uma dúvida de que a observação possa pertencer a um de dois clusters. Observações com um S_i negativo estão, provavelmente, no cluster errado.

Este processo de *clustering* requer *à priori* três escolhas importantes que se encontram listadas de seguida: i) escolha das variáveis a incluir no método CLARA; ii) escolha do número de *clusters* que se deseja obter; e, iii) escolha do tamanho da amostra que se quer utilizar para construir os *clusters*.

A escolha das variáveis a incluir é um processo efetuado pelo utilizador com base no seu conhecimento sobre os dados sendo que só após esta análise é possível escolher o número de *clusters* a aplicar. Neste caso, para escolher o número de *clusters* ideal foi utilizado o método da silhueta média, a qual permite verificar com que número de clusters se maximiza a silhueta média permitindo auferir a qualidade do processo de *clustering* através do processo descrito

²³ <http://www.sthda.com/english/wiki/print.php?id=241>

anteriormente. O tamanho da amostra baseia-se num processo iterativo que só acaba quando se obtiver um elevado coeficiente da silhueta.

Sendo o objetivo deste estudo perceber se os indicadores macroprudenciais estão relacionados com o rácio de crédito vencido, na análise de clusters a variável rácio de crédito vencido foi deixada de fora de forma a perceber se sem ela é possível agregar os contratos em 2 grupos, num os que apresentam crédito vencido e noutro, os que não apresentam.

O *software* utilizado para a análise de *clusters* foi o R (3.5.1) e os *inputs* necessários para utilizar o método CLARA estão descritos no Anexo IV.

3.6. ANÁLISE DE REGRESSÃO – REGRESSÃO DE DUAS PARTES

Muitas variáveis (y_i) em diversos tipos de estudo têm uma distribuição atípica no sentido em que são constituídas por valores contínuas positivas mas que no entanto incorporam muitos valores zero, ou seja, nos dados são observadas duas características: i) $y_i > 0$; ou ii) $y_i = 0$, onde a segunda característica é observada um conjunto de vezes suficientes de forma a necessitar de um tratamento especial. Variáveis com este tipo de distribuição são usualmente denominadas de variáveis censuradas, e neste caso, abaixo de zero.

Segundo Cameron e Trivedi (2005), pela natureza destas variáveis é normal que as mesmas não tenham uma média condicional linear em x , mesmo que a média da população seja linear, e, por isso uma estimação por OLS pode levar a uma estimação inconsistente do parâmetro relativo ao declive da reta, levando a que se obtenha um declive menos acentuado. Existem diversos modelos que permitem lidar com este tipo de variáveis, como por exemplo, o modelo censurado MLE e o modelo Tobit. No entanto, estes modelos não permitem que se utilize um modelo para estimar o facto de o acontecimento ser observado ou não e, um modelo diferente para estimar o acontecimento condicionado a este ser observado. Por exemplo, ao estudar as despesas hospitalares anuais dos clientes disponíveis na base de dados um primeiro processo determina se o paciente foi hospitalizado e o segundo processo explica as despesas que existiram com a hospitalização.

O modelo de duas partes permite modelar situações em que exista uma mistura de valores discretos e contínuos, acomodando o número massivo de zeros que os dados possam conter, o que por sua vez adiciona flexibilidade. Este modelo adequa-se a dados em que os zeros não correspondem a valores censurados mas sim a “verdadeiros” valores zero. Neste tipo de modelos é necessário a escolha de um modelo binário (*probit* ou *logit*), que é utilizado para modelar a probabilidade de ser observado um valor positivo comparativamente com a probabilidade ser observado um valor igual a zero. Após esta etapa e caso se observe um resultado positivo é necessário escolher um modelo de regressão apropriado para estimar este resultado positivo (OLS ou GLM). O modelo em questão não faz qualquer tipo de hipóteses sobre a correlação entre os erros das equações binária e contínua, ou seja, assume-se que os mecanismos das duas partes são independentes.

O modelo de duas partes permite que seja criada uma densidade diferente para os zeros e para os valores contínuos. Usualmente os zeros são modelados com base na probabilidade de se obter um resultado positivo.

Sendo um indivíduo que é totalmente observado denominado de participante na atividade em estudo e sendo também definida uma variável binária d em que $d = 1$ seja utilizada para participantes e $d = 0$ para não participantes. Suponha-se que $y > 0$ é observado para os participantes e $y = 0$ é observado para não participantes. Para os não participantes observa-se apenas $\Pr[d = 0]$. Para participantes a densidade condicional de y dado $y > 0$ é especificada como $f(y|d = 1)$, para uma escolha de densidade $f(\cdot)$. O modelo de duas partes para y é dado por:

$$f(y|\mathbf{x}) = \begin{cases} \Pr[d = 0|\mathbf{x}] & \text{se } y = 0 \\ \Pr[d = 1|\mathbf{x}] f(y|d = 1, \mathbf{x}) & \text{se } y > 0 \end{cases} \quad (4)$$

Em que \mathbf{x} corresponde ao vetor $(1 \times k)$ de variáveis explicativas.

Este modelo foi apresentado por Cragg (1971) como uma generalização do modelo de Tobit. O mais usual para modelar a decisão de participação d é ser utilizado um modelo probit ou logit. Neste caso a formulação é a seguinte:

$$F(y > 0) = \Pr(y > 0|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta} + u \quad (5)$$

Neste modelo \mathbf{x} corresponde ao vetor $(1 \times k)$ de variáveis explicativas, $\boldsymbol{\beta}$ o vetor $(k \times 1)$ de coeficientes da regressão e F representa uma função de distribuição cumulativa (*probit* ou *logit*).

Para valores positivos, usualmente é utilizado um modelo linear estimado por OLS, ou um modelo GLM. Este caso tipicamente é representado por:

$$G(y|y > 0, \mathbf{x}) = g(\mathbf{x}\boldsymbol{\gamma}) \quad (6)$$

Em que \mathbf{x} corresponde ao vetor de variáveis explicativas, $\boldsymbol{\gamma}$ representa o vetor de parâmetros a ser estimados e g representa a função de densidade apropriada (OLS ou GLM) para $y|y > 0$.

A média deste modelo pode ser escrita como o produto do valor esperado da primeira parte e da segunda parte do modelo:

$$E(y|\mathbf{x}) = \Pr(y > 0|\mathbf{x}) \times E(y|y > 0, \mathbf{x}) \quad (7)$$

Estimar os parâmetros deste modelo é muito simples pois a primeira parte $\Pr(y > 0|\mathbf{x})$ é modelada através de um modelo para variáveis binárias e para a segunda parte $E(y|\mathbf{x}, y > 0)$ é utilizado um modelo de resultados contínuos condicionado a ser observado um valor positivo, em que os erros das equações não têm de ser independentes para que se possa obter estimativas consistentes para as duas partes do modelo.

As estimativas de y_i , $(\hat{y}_i|\mathbf{x}_i)$ podem ser obtidos através da multiplicação dos valores estimados de cada parte do modelo, observação a observação, que equivale a

$$\hat{y}_i|x_i = (\hat{p}_i|x_i) \times (\hat{y}_i|y_i > 0, x_i) \quad (8)$$

Em que $\hat{p}_i|X_i$ corresponde ao valor estimado da probabilidade de $y_i > 0$.

Este modelo não tem como objetivo prever a probabilidade dos créditos entrarem em incumprimento, mas sim verificar a importância das variáveis independentes tanto para explicar a probabilidade do crédito estar em incumprimento, bem como para explicar o impacto das variáveis caso o crédito esteja em incumprimento e, por isso, o rácio de crédito vencido seja positivo.

O *software* utilizado para a análise da regressão de duas partes foi o STATA15.

3.6.1. Estimação do modelo

Sendo o objetivo desta etapa, por um lado verificar os impactos das diversas variáveis no rácio de crédito vencido, e por outro verificar quais as variáveis que mais impacto tiveram no mesmo para os diferentes períodos em análise, foi estimado apenas um modelo que engloba os três períodos em análise. Recorreu-se à criação de três variáveis binárias (uma para cada período) e foi decidido que a categoria de base seria a categoria relativa ao período da crise, o que tornará mais intuitiva a verificação de diferenças significativas entre o período da crise e o período anterior e posterior à mesma.

Como referido anteriormente, para esta etapa não foram utilizadas quatro das variáveis utilizadas para a análise de *clusters* e descritas no Anexo I (AnoCelebração, MntDivida, LTI e Incumprimento_2017) uma vez que estas variáveis apresentam níveis elevados de correlação com outras variáveis do modelo.

Segundo Deb e Norton (2018), para estimar a primeira parte do modelo normalmente é utilizado um modelo *probit* ou um modelo *logit* já que especificações alternativas para estimar a primeira parte do modelo geram resultados idênticos. Afirmam também que a diferença entre utilizar um *probit* ou um *logit* nunca é muito substancial. Por outro lado, a especificação do modelo utilizado para estimar a segunda parte é muito importante já que diferentes modelos podem levar a resultados muito diferentes. Para estimar a segunda parte do modelo existem diversas especificações que podem ser utilizadas, de forma a aliviar a potencial não-normalidade da variável explicativa. Muitos autores aplicam uma transformação logarítmica à variável dependente. Outros autores apontam a utilidade de se recorrer a um modelo GLM. Por exemplo Deb e Norton (2018), afirmam que os modelos GLM permitem acomodar de forma natural a assimetria dos dados, dando ao analista uma maior flexibilidade. Nesta estimação da segunda parte do modelo pode, então, ser utilizado um modelo linear estimado por OLS simples ou aplicando uma transformação logarítmica à variável dependente para aplicar também um modelo estimado por OLS ou um modelo GLM que requer que seja escolhida a família da distribuição e uma função de ligação que relaciona o valor esperado da variável independente com as covariâncias e com os coeficientes estimados. Esta função de ligação define a relação das x variáveis diretamente com a média da distribuição de y em vez de modificar a distribuição das variáveis observadas.

De forma a determinar quais das variáveis são mais importantes para explicar o rácio de crédito vencido, as variáveis independentes foram estandardizadas. Desta forma as variáveis com um coeficiente absoluto mais elevado serão as mais importantes para explicar o rácio de crédito vencido. Segundo Kim e Dong-ku (2011), estandardizar as variáveis implica remover a média de cada valor observado e dividir pelo desvio-padrão da variável permitindo que as variáveis dependentes sejam interpretadas da maneira usual. Segundo estes autores estandardizar as variáveis dependentes traz também benefícios grandes quando o modelo inclui termos de interação ou polinomiais pois estes termos produzem multicolineariedade.

3.6.2. Avaliação dos modelos

Testes de hipóteses

Após obter as estimativas para os coeficientes iremos avaliar se são significativos, ou seja, verificar se as variáveis utilizadas no modelo estão significativamente associadas à variável dependente. Para esta avaliação recorreu-se a um teste de significância individual.

Significância individual

Os testes de hipótese individuais para os coeficientes da regressão são essenciais para determinar se, individualmente, as variáveis explicativas são importantes no modelo de regressão para que seja possível identificar variáveis independentes que não expliquem a variável dependente.

As hipóteses deste teste são dadas por,

$$\begin{cases} H_0: \beta_j = 0 \\ H_1: \beta_j \neq 0 \end{cases}; j = 0, 1, \dots, k \quad (9)$$

Em que k representa o número de parâmetros do modelo.

Com base na tabela 6, é possível verificar que todas as variáveis incluídas no modelo são estatisticamente significativas a um nível de significância de 10%. Assim, rejeitamos H_0 com um nível de significância de 10%, o que quer dizer que, individualmente, as variáveis são relevantes para o modelo.

3.6.3. Qualidade dos modelos

Critério de Informação

Duas outras medidas, também utilizadas para avaliar a qualidade do modelo são o critério de informação de Akaike (AIC) e o critério de informação Bayesiano (BIC). Estas estatísticas têm por base o logaritmo da verosimilhança, penalizando um modelo que tenha muitas variáveis.

A medida AIC é dada por:

$$AIC = -2 [\text{Log}(L) - k] \quad (10)$$

Em que k representa o número de parâmetros do modelo e L o valor da verossimilhança para o modelo estimado.

Enquanto a medida BIC é dada por:

$$BIC = -2 \text{Log}(L) + \text{Log}(n) \times k \quad (11)$$

Em que k representa o número de parâmetros do modelo, L o valor da verossimilhança para o modelo estimado e N representa o número de observações do modelo.

Ambas as medidas permitem comparar modelos quer estes sejam aninhados ou não. Estes critérios penalizam modelos com muitas variáveis, e por isso um menor valor para esta medida significa que menos informação será perdida.

3.6.4. Seleção de modelos

A fase de seleção de modelos é uma etapa bastante importante na investigação. Deve ser escolhido um modelo simples mas que descreva o melhor possível os dados observados. Geralmente existem muitas variáveis que podem ser relevantes para explicar a variabilidade da variável dependente, o que acaba por tornar o processo de seleção das variáveis e de diversas combinações entre as mesmas um processo difícil e demorado.

Segundo Deb e Manning (2015), devido às dificuldades de lidar com heterocedasticidade, deve utilizar-se um modelo GLM com uma função de ligação para que seja possível obter estimadores e efeitos marginais consistentes, pois ao utilizar um modelo linear estimado por OLS, mesmo que transformado utilizando uma abordagem não paramétrica como a de Duan (1893), é muito comum que exista heterocedasticidade nos erros do modelo.

Uma vez que a literatura recomenda a utilização de um modelo GLM em vez de um modelo OLS na estimação da segunda parte do modelo, optou-se por utilizar duas combinações de modelos diferentes para verificar qual destes permitiria descrever o melhor possível os dados observados. Os modelos resultam da conjugação de um modelo *probit* ou *logit* para estimar a primeira parte e de um tipo de regressão da classe de modelos GLM (gama), utilizando um estimador robusto a más especificações do modelo. Optou-se por utilizar as medidas AIC e BIC para verificar qual dos modelos utilizados apresenta um melhor nível de ajustamento.

Tabela 5 – Comparação entre os modelos

Modelo	Método	AIC	BIC
1	<i>Logit + Gama</i>	62 043,57	62 725,74
2	<i>Probit + Gama</i>	61 621,5	62 303,67

A partir da tabela anterior é possível verificar que o modelo 2, que utiliza um *probit* para estimar a primeira parte e um modelo GLM (utilizando a distribuição de gama) para estimar a segunda parte do modelo é o modelo que apresenta um melhor ajustamento pois apresenta um menor valor para

as medidas AIC e BIC, medidas estas que medem o grau de informação perdida com o ajuste do modelo.

As variáveis independentes incluídas em ambas as partes do modelo são os seguintes:

- MntInicial;
- RendAnualBrutoDataIni;
- PrecoAqHipotecado;
- ValorPrimeiraAval;
- TAN;
- Maturidade;
- ValorPrimeiraPrest;
- O_SII;
- DifAnocelAnoref;
- DSTI;
- LTV;
- MntInicial x AC;
- RendAnualBrutoDataIni x AC;
- PrecoAqHipotecado x AC;
- ValorPrimeiraAval x AC;
- TAN x AC;
- Maturidade x AC;
- ValorPrimeiraPrest x AC;
- O_SII x AC;
- DifAnocelAnoref x AC;
- DSTI x AC;
- LTV x AC;
- MntInicial x DC;
- RendAnualBrutoDataIni x DC;
- PrecoAqHipotecado x DC;
- ValorPrimeiraAval x DC;
- TAN x DC;
- Maturidade x DC;
- ValorPrimeiraPrest x DC;
- O_SII x DC;
- DifAnocelAnoref x DC;
- DSTI x DC;
- LTV x DC;

4. RESULTADOS E DISCUSSÃO

4.1. ANÁLISE DE CLUSTERS

Nesta secção irão ser apresentados os resultados da análise de *clusters*. Os *clusters* foram utilizados de forma a verificar se observações com crédito vencido seriam agrupadas no mesmo cluster, deixando de fora a análise, a variável rácio de crédito vencido e por isso verificar se observações com rácio de crédito vencido têm comportamentos semelhantes nas variáveis em análise.

4.1.1. Antes da Crise

Uma vez que o período antes da crise é o período com um maior volume de dados o R não consegue calcular o método da silhueta média, logo não é possível perceber qual o número de clusters que maximiza a silhueta média e por isso foram testadas várias soluções, com diferentes números de clusters definidos à priori ($k=2,3,4$) e testados diferentes tamanhos para a amostra (samples no intervalo de]50;8000[). Desta forma, a escolha do número ideal de *clusters* foi feita com o k e o número de samples que maximizava o valor da silhueta média. O melhor resultado resulta de 2 *clusters*, formados com o método CLARA, baseados em 500 amostras derivadas do *dataset*.

Analisando o resultado da análise de *clusters* é possível verificar que, mesmo sem ser utilizada a variável rácio de crédito vencido à análise, todas as observações que apresentam crédito vencido foram agrupadas no mesmo cluster (*cluster 2*).

A silhueta do melhor resultado obtido está representada na figura 9. A silhueta representa o quão bem cada observação se assemelha ao *cluster* a qual foi atribuído, isto é feito comparando o quão perto os objetos estão do seu próprio cluster comparativamente com os objetos no outro cluster.

A silhueta da figura 9, que utiliza as duas primeiras componentes principais para fazer o gráfico de pontos, apresenta um coeficiente médio de 0,72, o que significa que existe uma forte estrutura nos dados, existindo apenas algumas observações no *cluster 1* que apresentam um valor inferior. Uma vez que não existem observações perto de zero e nenhuma das observações apresenta valores negativos o resultado aparenta ser adequado.

Adicionalmente é possível verificar, na figura 10, que as duas primeiras componentes principais do conjunto de dados utilizados para formar os *clusters* permitem explicar cerca de 44% da variabilidade dos dados.

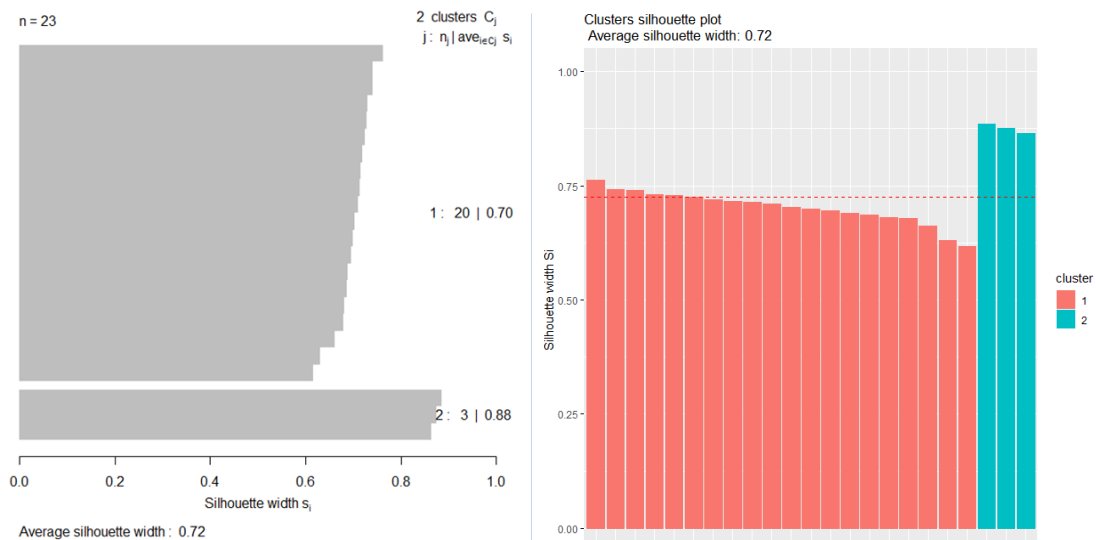


Figura 10 - Silhueta do melhor resultado obtido com o método CLARA

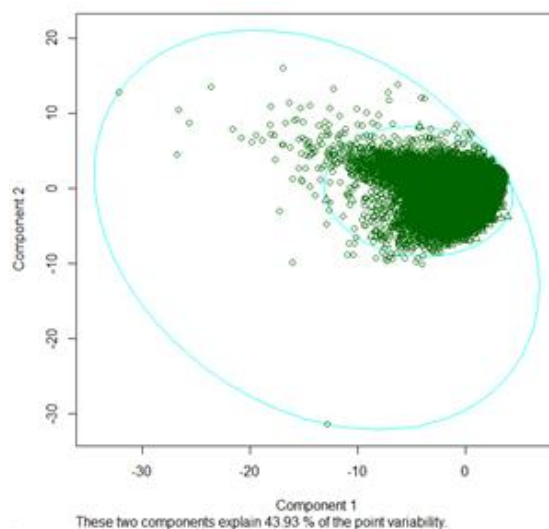


Figura 11 - Gráfico de *clusters* utilizando o melhor resultado obtido com o método CLARA

Para melhorar a apresentação visual dos dois *clusters* formados, é apresentado na tabela seguinte, os valores médios das variáveis mais relevantes para o estudo. A verde estão os valores superiores ou iguais à média do respetivo indicador e a vermelho os valores inferiores à média.

Tabela 6 - Valores médios dos indicadores para o período de 2005 a 2009

Cluster	Ano de Celebração	Montante Inicial	Montante Dívida	Rendimento Inicial	Preço do bem hipotecado	Valor 1ª Avaliação	TAN	Maturidade
1	2 006,27	91 996,80	66 286,17	28 647,79	108 476,75	162 450,2	0,51	35,36
2	2 006,24	87 772,41	70 226,73	26 674,73	99 763,67	149 029,3	1,33	36,82
								Quantidade de contratos no cluster
Cluster	Valor 1ª Prestação	O-SII	DSTI	LTI	LTV	RCV		
1	432,27	0,52	30,08	5,41	91,52	0,00		72 015
2	411,96	0,51	31,84	5,80	92,47	0,79		2 650

Podemos então verificar que temos 2 grupos de agregados de contratos distintos:

- **Cluster 1 (72 015 observações):** Contratos que não se encontram em incumprimento e por isso não têm RCV pois tinham características de concessão de crédito desejáveis (LTI, DSTI, LTV, Maturidade e TAN abaixo da média).
- **Cluster 2 (2 650 observações):** Contratos que se encontram em incumprimento com um RCV acima da média por terem sido concedidos com características de concessão de crédito não tão desejáveis (LTI, DSTI, LTV, Maturidade e TAN acima da média), ou seja, contratos que foram concedidos com um maior risco por parte das instituições.

4.1.2. Durante a crise

Para o período da crise já foi possível utilizar o método da silhueta média para calcular o número ótimo de *clusters*. Utilizando o método da silhueta média, é possível verificar no Anexo VII que o número ótimo de clusters é dois.

Testando também diferentes tamanhos para a amostra, o melhor resultado resulta de 2 grupos de *clusters*, formados com o método CLARA, baseados em 1000 amostras derivadas do dataset.

Com as duas primeiras componentes principais que resultam dados utilizados para formar os *clusters* no período da crise é possível explicar cerca de 42% da variabilidade dos dados.

A silhueta do melhor resultado obtido está representada na figura 11, esta apresenta um coeficiente médio de 0,82, significando que existe uma forte estrutura nos dados, existindo poucas observações no *cluster* 1 que apresentam um valor inferior.

Ao contrário do verificado no período antes da crise, os *clusters* formados para o período da crise já não são tão bem definidos. Aproximadamente 1,5% dos contratos foram alocados ao *cluster* contrário: i) 1,2% dos contratos com crédito vencido foram alocados ao *cluster* dos contratos sem crédito vencido, o que significa que apesar do crédito apresentar boas condições de concessão de crédito o mutuário não conseguiu pagar o seu crédito; ii) 0,3% dos créditos sem crédito vencido foram alocados ao *cluster* dos contratos com crédito vencido, significando que estes créditos não aparentavam ter boas condições de concessões de crédito, no entanto o mutuário ainda está capaz de pagar o crédito em questão.

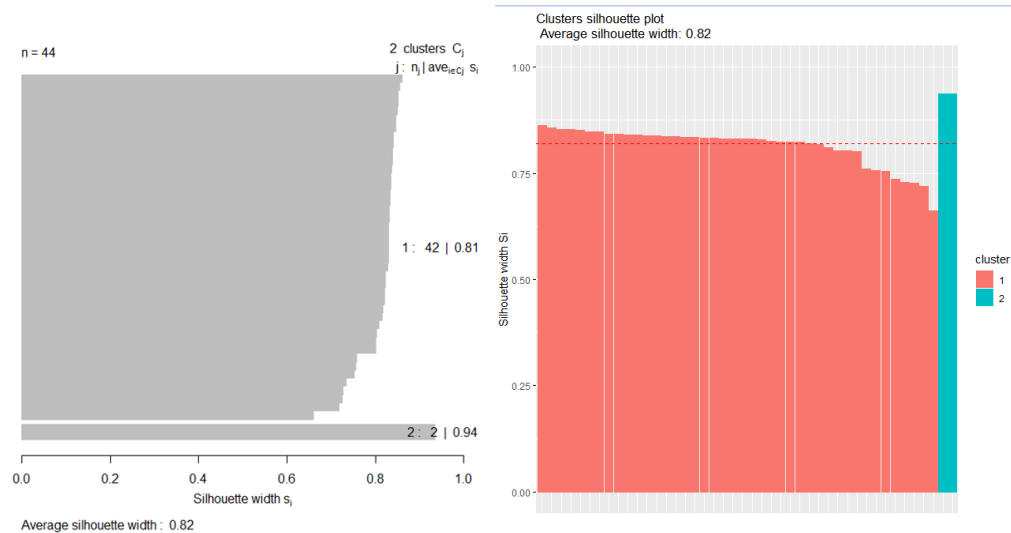


Figura 12 - Silhueta do melhor resultado obtido com o método CLARA

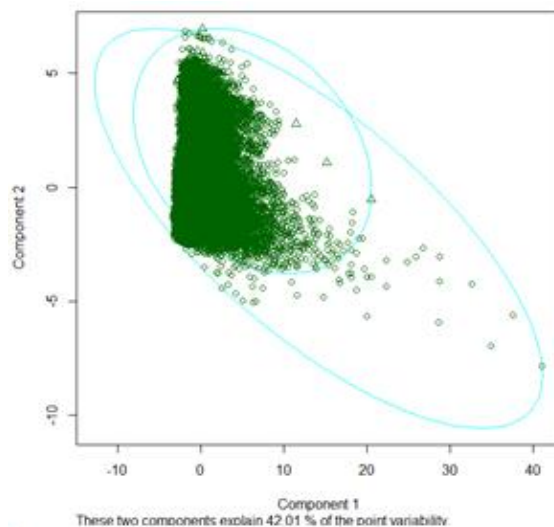


Figura 13 - Gráfico de *clusters* utilizando o melhor resultado obtido com o método CLARA

Para o período da crise os valores médios das variáveis mais importantes para o estudo encontram-se resumidas na figura seguinte.

Tabela 7 - Valores médios dos indicadores para o período de 2010 a 2014

Cluster	Ano de Celebração	Montante Inicial	Montante em Dívida	Rendimento Inicial	Preço do bem hipotecado	Valor 1ª Avaliação	TAN	Maturidade
1	2 009,23	95 339,85	75 649,64	32 844,64	114 329,0	167 142,8	0,95	35,28
2	2 009,00	91 226,57	75 945,89	24 338,73	107 854,8	140 615,9	1,89	35,72
Cluster	Valor 1ª Prestação	O-SII	DSTI	LTI	LTV	RCV	Quantidade de contratos no cluster	
1	423,80	0,53	25,56	4,70	93,67	0,01	64 449	
2	476,42	0,89	121,48	28,37	94,55	0,70	1 190	

Os *clusters* formados para o período da crise apresentam as seguintes características:

- **Cluster 1 (64 449 observações):** Este cluster engloba poucos contratos que se encontram em incumprimento (776 contratos) e por isso apresenta um valor baixo para a variável RCV. Os contratos deste *cluster* têm boas características de concessão de crédito (LTI, DSTI, LTV, Maturidade e TAN abaixo da média).
- **Cluster 2 (1 190 observações):** Contratos que apresentam um risco maior para as instituições, pois estes não têm tão boas características de crédito (LTI, DSTI, LTV, Maturidade e TAN acima da média) e por isso apresentam um rácio de crédito vencido acima da média.

4.1.3. Depois da Crise

Finalmente, para o período posterior à crise foi também possível calcular o número ótimo de *clusters* através do método da silhueta média apresentado no Anexo VII. Esta análise indicou que dois *clusters* seria o número de *clusters* que maximizava os resultados.

O melhor resultado obtido deriva de 2 grupos de *clusters* formados com o método CLARA, baseados em 1000 amostras com 150 observações derivadas do *dataset*.

Para o período depois da crise, a análise de *clusters* também agrupa perfeitamente os contratos com crédito vencido num *cluster* (*cluster* 2) e todos os contratos sem crédito vencido num outro *cluster*.

A silhueta do melhor resultado obtido está representada na figura 13. Esta apresenta um coeficiente médio de 0,97, o que significa que existe uma estrutura muito forte nos dados, existindo apenas algumas observações no primeiro *cluster* que apresentam um valor inferior, no entanto o resultado aparenta ser adequado. Com as duas primeiras componentes principais que resultam dados utilizados para formar os *clusters* no período da crise é possível explicar cerca de 42% da variabilidade presente nos dados.

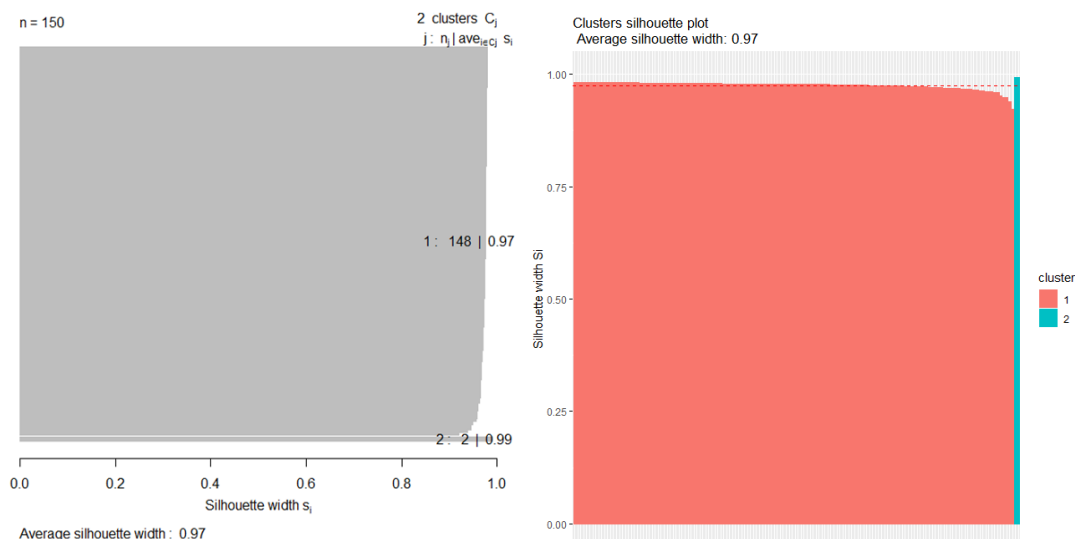


Figura 14 - Silhueta do melhor resultado obtido com o método CLARA

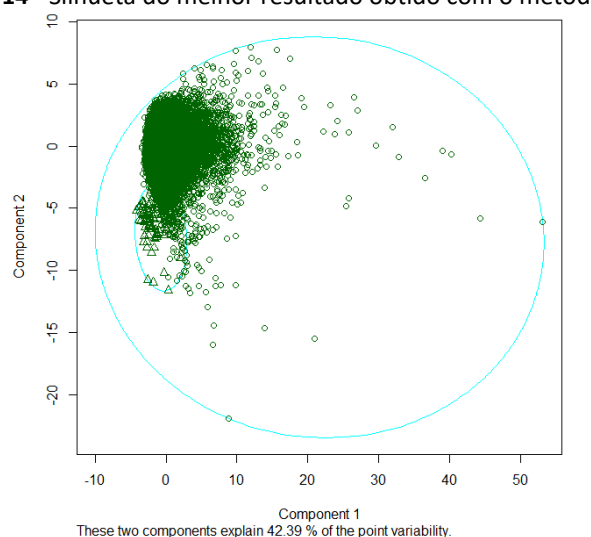


Figura 15 - Gráfico de *clusters* utilizando o melhor resultado obtido com o método CLARA

Na figura 16 estão representados os valores médios das variáveis mais relevantes que derivam dos 2 *clusters* criados com o método CLARA.

Tabela 8 - Valores médios dos indicadores para o período de 2015 a 2017

Cluster	Ano de Celebração	Montante Inicial	Montante em Dívida	Rendimento Inicial	Preço do bem hipotecado	Valor 1ª Avaliação	TAN
1	2 016,27	108 450,89	103 223,28	35 268,20	140 637,52	169 892,30	1,85
2	2 015,56	89 414,75	80 786,78	20 784,88	95 900,73	121 961,1	3,78

Cluster	Maturidade	Valor 1ª Prestação	O-SII	DSTI	LTI	LTV	RCV	Quantidade de contratos no cluster
1	35,88	391,60	1,00	20,65	4,69	84,00	0,00	27647
2	37,51	366,24	0,00	29,91	5,90	92,92	0,97	80

Podemos então verificar que temos 2 grupos de agregados de contratos distintos:

- **Cluster 1 (27 647 observações):** Contratos que não se encontram em incumprimento e por isso não têm crédito vencido já que tinham características de concessão de crédito desejáveis (LTI, DSTI, LTV e TAN abaixo da média). Todos os créditos do *cluster 1* parecem ter sido celebrados apenas com instituições classificadas como de grande importância sistémica.
- **Cluster 2 (80 observações):** Contratos que se encontram em incumprimento com um crédito vencido acima da média por terem sido concedidos com características de concessão de crédito menos desejáveis (LTI, DSTI, LTV e TAN acima da média). Ao contrário do *cluster 1*, neste cluster apenas existem créditos celebrados com instituições que não classificadas como de grande importância sistémica.

4.2. ANÁLISE DE REGRESSÃO

4.2.1. Estimação dos modelos

Nesta secção irá ser apresentado o resultado da estimação do modelo de duas partes. Foi utilizado o modelo de duas partes para explicar o rácio de crédito vencido dos mutuários, tendo em conta o número massivo de zeros que esta variável apresenta.

O modelo de duas partes escolhido para estimar o impacto das diferentes variáveis no rácio de crédito vencido utiliza um *probit* para modelar a probabilidade de se observar um valor positivo no rácio de crédito vencido comparativamente com a probabilidade de ser observado um valor igual a zero e um modelo GLM (utilizando a distribuição gama) para estimar o rácio de crédito vencido condicionado a efetivamente ser observado um valor positivo.

4.2.2. Interpretação dos coeficientes de regressão

Após a seleção do modelo e da sua avaliação é necessário interpretar os valores associados aos coeficientes do modelo escolhido. Esta interpretação depende da natureza das variáveis explicativas que podem ser contínuas ou binárias.

Os regressores ValorPrimeiraAval, ValorPrimeiraPrest foram removidos do modelo por não serem estatisticamente significantes no modelo combinado (a um nível de 10%). Adicionalmente, os regressores RendAnualBrutoDataIni_AC, PrecoAqHipotecado_AC, LTV_AC, MntInicial_DC, ValorPrimeiraAval_AC, ValorPrimeiraPrest_AC, TAN_DC, ValorPrimeiraAval_DC, Maturidade_DC, RendAnualBrutoDataIni_DC, PrecoAqHipotecado_DC, ValorPrimeiraPrest_DC, O_SII_DC, que derivam de interações com as variáveis relativas ao período da crise não aparentam ser estatisticamente significativos (a um nível de 10%) e por isso foram também removidos do modelo inicial.

Removendo os regressores acima descritos é possível observar, na tabela 9, que todos os restantes regressores são estatisticamente significativos (a um nível de 10%) no modelo combinado.

A tabela 9 apresenta os resultados da estimativa dos coeficientes de regressão, erro-padrão e a significância para cada variável do modelo selecionado. As estimativas dos coeficientes permitem quantificar o efeito de cada uma das variáveis explicativas no rácio de crédito vencido. O erro-padrão dos estimadores permite avaliar a significância dos coeficientes estimados. Neste sentido, compara-se o valor do teste de significância individual com o quantil de probabilidade da normal padrão, que depende do nível de significância que se pretende.

Todas as variáveis afetam o rácio de crédito vencido, positivamente ou negativamente. É possível verificar que, no modelo combinado, as variáveis que não contemplam interações têm os sinais esperados e, estes encontram-se descritos no Anexo VIII.

Tabela 9 – Resultado da estimação do modelo de duas partes

Variáveis explicativas	Efeitos parciais		
	Probit	GLM (Gama)	Combinado
MntInicial	0,051*** (0,020)	0,009* (0,006)	0,002*** (0,001)
RendAnualBrutoDataIni	0,040*** (0,007)	-0,008** (0,003)	0,002*** (0,0003)
PrecoAqHipotecado	-0,121*** (0,021)	0,00001 (0,005)	-0,005*** (0,001)
TAN	0,272*** (0,007)	0,027*** (0,003)	0,013*** (0,0004)
Maturidade	0,079*** (0,010)	0,115*** (0,006)	0,006*** (0,0005)
O_SII	-0,059*** (0,012)	0,002 (0,003)	-0,003*** (0,001)
DifAnocelAnoref	0,625*** (0,029)	-0,053*** (0,010)	0,027*** (0,001)
DSTI	0,003 (0,002)	0,013 (0,008)	0,0004** (0,0002)
LTV	0,019*** (0,007)	0,001 (0,002)	0,001*** (0,0003)
MntInicial_AC	0,029* (0,016)	0,001 (0,005)	0,001* (0,001)
TAN_AC	0,066*** (0,007)	-0,011*** (0,003)	0,003*** (0,0003)
Maturidade_AC	0,205*** (0,032)	0,058*** (0,016)	0,010*** (0,001)
O_SII_AC	-0,092*** (0,014)	0,002 (0,004)	-0,004*** (0,001)
DifAnocelAnoref_AC	-0,367*** (0,035)	-0,054*** (0,018)	-0,017*** (0,002)
DSTI_AC	0,031*** (0,009)	0,002 (0,003)	0,001*** (0,0004)
DifAnoDCelAnoref_DC	0,037* (0,019)	0,024*** (0,006)	0,002** (0,001)
DSTI_DC	0,079*** (0,015)	-0,002 (0,005)	0,003*** (0,001)
LTV_DC	-0,082** (0,033)	-0,029** (0,012)	-0,004*** (0,001)
Constante	-2,156*** (0,010)	0,799*** (0,005)	0,021*** (0,0003)
Nº de observações	168 031	4 474	
Pseudo R ²	11,09%	-	
Log pseudolikelihood	-	-3 496,05	

Notas:

1. O erro-padrão de cada estimativa encontra-se dentro de parênteses.
2. Níveis de significância *10%, **5%, ***1%

Efeitos parciais do modelo combinado

Analisando os resultados do modelo de duas partes apresentados na tabela 9 é possível verificar que os efeitos parciais para o montante inicial, para a TAN, para a maturidade e, por fim, para o DSTI e para o LTV aplicados ao contrato de crédito apresentam valores positivos em ambas as partes do modelo, sendo estas variáveis também estatisticamente significativas ao nível de significância de 1%. As variáveis relacionadas com o rendimento anual bruto do mutuário, com o preço de aquisição do imóvel, o facto de o crédito ser contratado numa instituição classificada como de risco sistémico e, por fim, o ano em que o contrato foi celebrado apresentam sinais diferentes nas duas partes do modelo, sendo estas também estatisticamente significativas ao nível de significância de 1%.

Todas as variáveis do modelo que contêm termos de interação aparentam ter diferenças bastante significativas relativamente ao período da crise. É possível verificar que existem diferenças bastante significativas na TAN, no montante inicial, na maturidade e no facto de o contrato ser contratado numa instituição classificada como de risco sistémico entre o período da crise e o período antes da crise, no entanto, esta diferença não aparenta ser significativa comparativamente com o período pós-crise.

É também possível de verificar que as variáveis que aparentam ter um impacto mais significativo no rácio de crédito vencido são: i) o número de anos entre a celebração do contrato e ano em que foi reportado o contrato, 2017 (o efeito marginal do ano de celebração tem um efeito médio de 0,027p.p no rácio de crédito vencido); ii) a TAN do contrato (o efeito marginal da TAN tem um efeito médio de 0,013p.p no rácio de crédito vencido); iii) a maturidade do contrato (o efeito marginal da maturidade do contrato tem um efeito médio de 0,006p.p no rácio de crédito vencido); iv) o preço de aquisição do imóvel (o efeito marginal da do preço de aquisição do imóvel tem um efeito médio de -0,005p.p no rácio de crédito vencido); e, por fim v) o facto de o crédito ser contratado numa instituição classificada como de risco sistémico (o efeito marginal do contrato ser celebrado numa instituição considerada como de risco sistémico tem um efeito médio de -0,003p.p no rácio de crédito vencido).

Relativamente às variáveis macroprudenciais em estudo, para além da maturidade, que demonstra ter um impacto importante no rácio de crédito vencido, podemos verificar que ainda que as restantes (LTV e DSTI) tenham um impacto inferior na variável de interesse, é possível verificar que o efeito marginal do LTV tem um efeito médio de 0,001p.p no rácio de crédito vencido e o DSTI um efeito médio de 0,0004p.p., mantendo tudo o resto constante. É também possível verificar que o LTV praticado no período pós-crise tem um impacto inferior ao LTV praticado no período da crise tanto relativamente à probabilidade do crédito se encontrar em incumprimento como no rácio de crédito vencido, condicionado ao crédito se encontrar em incumprimento. O mesmo já não se verifica comparando o LTV praticado no período antes da crise, já que esta variável apresentava um nível de significância superior a 10%. Relativamente ao DSTI é possível verificar diferenças significativas entre o período da crise e o período antes da crise e o período pós-crise, sendo possível verificar que: i) o DSTI praticado no período antes da crise tem um impacto superior ao DSTI

praticado no período da crise tanto relativamente à probabilidade do crédito se encontrar em incumprimento como no rácio de crédito vencido, condicionado ao crédito se encontrar em incumprimento; e, ii) o DSTI praticado no período antes da crise tem um impacto superior ao DSTI praticado no período da crise relativamente à probabilidade do crédito se encontrar em incumprimento, no entanto, estando o crédito condicionado a estar em incumprimento este apresenta um DSTI inferior no período pós-crise comparativamente com o período da crise. Estes resultados demonstram que tanto a instituição que concede o crédito como os mutuários beneficiam de contratos com DSTI e LTV mais reduzidos de forma a reduzir o crédito em incumprimento. Adicionalmente, através das estatísticas descritivas apresentadas no Anexo V, é possível verificar que no período pós crise as instituições passaram a ser mais cautelosas nos seus critérios de concessão de crédito, concedendo crédito com valores de DSTI e de LTV inferiores aos anteriormente praticados, o que por sua vez também contribuirá para níveis de incumprimento mais baixos.

Efeitos parciais do modelo *probit* e do modelo GLM

O modelo *probit* foi utilizado para estimar a probabilidade de um contrato estar em incumprimento e, por isso, apresentar um valor diferente de zero na variável de interesse comparativamente com a probabilidade do contrato não estar em incumprimento. Já o modelo GLM foi utilizado para estimar os impactos das diferentes variáveis em estudo, condicionado ao contrato estar em incumprimento. Para a primeira parte do modelo foram utilizadas as observações totais presentes na base de dados (168 031), no entanto para estimar a segunda parte do modelo foram apenas utilizadas cerca de 3% das observações que correspondem aos contratos em incumprimento em 2017.

As variáveis que aparentam ser mais relevantes para o modelo *probit* e para o modelo GLM são diferentes. Para o modelo *probit* as mais relevantes são: i) o número de anos entre a celebração do contrato e ano em que foi reportado o contrato, 2017; ii) a TAN do contrato; iii) o preço de aquisição do imóvel; iv) a maturidade do contrato; e, por fim v) o facto de o crédito ser contratado numa instituição classificada como de risco sistémico. Já para o modelo GLM as variáveis mais relevantes são: i) a maturidade do contrato; ii) o número de anos entre a celebração do contrato e ano em que foi reportado o contrato, 2017; iii) a TAN do contrato; iv) o DSTI aplicado ao contrato; e, por fim v) o montante inicial do contrato. Ou seja existem três variáveis (o número de anos entre a celebração do contrato e ano em que foi reportado o contrato, 2017, a TAN e a maturidade do contrato) que são bastante importantes tanto para a probabilidade do contrato estar em incumprimento como para a o impacto no rácio de crédito vencido de contratos que estão em incumprimento. Já o o preço de aquisição e o facto de o crédito ser contratado numa instituição classificada como de risco sistémico parecem apenas ser mais relevantes para a probabilidade do contrato estar ou não em incumprimento e não parecem ter tanto impacto no rácio de crédito vencido de contratos que se encontram em incumprimento. Pelo contrário, as variáveis relativas ao DSTI aplicado e ao montante inicial do contrato parecem não ser muito relevantes para a probabilidade do contrato entrar em

incumprimento, no entanto têm um impacto bastante grande no rácio de crédito vencido de contratos em incumprimento.

Avaliando agora os efeitos parciais do modelo *probit* e do modelo GLM é possível verificar a probabilidade do contrato estar em incumprimento aumenta, em cerca de 0,051p.p, 0,272p.p, 0,079p.p, 0,003p.p e 0,019p.p, respetivamente, com o aumento das variáveis montante inicial, TAN, maturidade, DSTI e LTV e, estando o crédito em incumprimento, o mesmo aumento nas variáveis descritas provoca um aumento, em cerca de 0,009p.p, 0,027p.p, 0,0115p.p, 0,013p.p e 0,001p.p, respetivamente, na média do rácio de crédito vencido.

Os créditos concedidos em instituições classificadas como de risco sistémico têm uma probabilidade inferior, em cerca de 0,06p.p de se encontrarem em incumprimento, no entanto, estando o crédito em incumprimento apresentam um rácio de crédito vencido superior, em cerca de 0,002p.p, em relação às instituições que não recaem nesta classificação. A probabilidade do crédito estar em incumprimento aumenta, cerca de 0,04p.p, com o aumento do rendimento bruto anual do mutuário, no entanto, estando o crédito em incumprimento este mesmo aumento no rendimento contribui, em média, para uma redução, em cerca de 0,008p.p, no rácio de crédito vencido. Já o oposto acontece quando analisamos a variável que diz respeito ao preço de aquisição do imóvel, já que a probabilidade do crédito estar em incumprimento diminui, em cerca de 0,121p.p, com o aumento do preço de aquisição do imóvel e, estando o crédito em incumprimento, o mesmo aumento no preço de aquisição provoca um aumento médio de 0,00001p.p no rácio de crédito vencido.

Por fim é possível verificar que o valor estimado do rácio de crédito vencido é de 0,0213 o que não está muito longe da média verdadeira de 0,0215.

5. CONCLUSÕES

Todo o processo de análise e avaliação de risco de crédito bancário e da análise da solvabilidade do mutuário adquire, nos dias de hoje, uma importância indubitável para o sucesso da atividade bancária. Assim, é necessário que as instituições disponham de certos indicadores e de um processo de análise de qualidade aquando da captação de um novo crédito.

A mais recente crise financeira global colocou os bancos sob uma atenta vigilância, o que se refletiu em critérios de concessão de crédito mais restritivos. O que leva ao ponto de partida deste estudo, verificar se o rácio de crédito vencido aumenta quando existem fatores de concessão de crédito não favoráveis para o mutuário. Tendo este objetivo em mente, foi concebido um modelo para explicar o rácio de crédito vencido onde foram considerados apenas fatores mensuráveis, divisíveis em cinco dimensões.

A primeira etapa deste processo que incidiu sobre a análise de *clusters* de grande dimensão (Cluster CLARA) auxiliou no agrupamento de observações com características semelhantes em *clusters* de contratos. Esta etapa permitiu concluir que, mesmo não utilizando a variável que diz respeito ao crédito vencido, foi possível alocar unicamente a um *cluster* os contratos com crédito vencido nos períodos antes e depois da crise. Para o período da crise já se verificou um erro de 2%, o que mesmo assim não é significativo. Este erro permitiu concluir que certos contratos mesmo apresentando boas características de concessão de crédito entraram em incumprimento e que certos contratos com características de concessão de crédito menos favoráveis não acabassem entrando em incumprimento.

A segunda etapa deste processo, que diz respeito à análise de regressão, utilizando um modelo de duas partes utiliza um modelo para modelar a primeira parte do modelo (*probit*), que permite estimar a probabilidade de um contrato se encontrar em incumprimento, e um modelo para modelar o impacto das variáveis independentes no rácio de crédito vencido (GLM com distribuição gama), condicionado à observação de um valor positivo.

Com base nos critérios de informação de AIC e BIC percebeu-se que o melhor modelo para estimar o rácio de crédito vencido seria um modelo *probit* acompanhado de um GLM com uma distribuição *gama*. Todas as variáveis incluídas no modelo são estatisticamente significativas no modelo conjunto, embora nem sempre sejam estatisticamente significativas nos modelos individuais, no entanto, pela razoabilidade económica foram mantidas nos modelos individuais.

De um modo geral, as variáveis que têm um maior impacto sobre a probabilidade do crédito estar em incumprimento e no rácio de crédito vencido são a TAN do contrato, o facto de o crédito ser contratado numa instituição classificada como de risco sistémico, o ano em que o contrato foi celebrado e, por fim, a maturidade. O LTV e o DSTI, que não tiveram um impacto tão forte no rácio de crédito vencido, também demonstram ter um impacto positivo na variável em estudo. Estes resultados permitem chegar ao mesmo tipo de conclusões retiradas da análise de *clusters*.

Sendo que a TAN do contrato engloba a taxa de referência e o *spread* seria de esperar que as instituições atribuísem um *spread* mais elevado aos clientes com maior risco, antecipando já o facto de estes contratos serem mais propícios a entrar em incumprimento. Esta prática possivelmente poderá explicar a relação positiva entre a TAN e o rácio de crédito vencido.

O facto de o crédito ser contratado numa instituição classificada como de risco sistémico ter uma relação negativa com a variável dependente pode, possivelmente, ser justificada pelo facto de serem as grandes instituições que mais contratos realizam por ano, que têm modelos de previsão de risco mais sofisticados e, por isso, têm uma maior capacidade para rejeitar contratos com um maior risco deixando estes contratos para as instituições de menor dimensão.

À partida não é expectável que um contrato entre em incumprimento poucos meses depois da sua celebração, já que não se espera que as condições financeiras dos mutuários se alterem em poucos meses, por esta razão é necessário algum tempo para que o contrato possa ou não entrar em incumprimento. Este fenómeno potencialmente justifica a relação positiva entre o ano de celebração do contrato e o rácio de crédito vencido.

As razões pela qual os mutuários sentem necessidade de estender o prazo de maturidade do seu contrato são variadas, no entanto a mais relevante, que possivelmente explica a relação positiva entre a maturidade e o rácio de crédito vencido, será o facto de estes mutuários não terem riqueza suficiente para fazer face às despesas associadas ao novo crédito e, eventualmente, aos créditos anteriores. Sendo assim, estes mutuários necessitam de estender a maturidade para que o valor da sua prestação seja suavizada ao longo da duração do contrato.

Por último sendo o DSTI um indicador que permite verificar a taxa de esforço do mutuário para fazer face ao novo empréstimo, e o LTV um indicador que fornece informação sobre as potenciais perdas para o sistema bancário em caso de incumprimento do crédito, monitorizando o comportamento dos bancos face ao risco no momento de concessão de crédito é de esperar que um aumento em qualquer um dos indicadores provoque um aumento no rácio de crédito vencido.

Assim, os resultados deste estudo permitem perceber que os indicadores macroprudenciais têm um impacto positivo no rácio de crédito vencido e, por isso, é relevante que sejam impostos limites para que a concessão de crédito seja baseada em critérios mais sustentáveis, tanto para os mutuários, como para o sistema financeiro para que a longo prazo se registem cada vez mais exista uma maior capacidade para cumprir com o crédito contratado.

5.1. RECOMENDAÇÕES FUTURAS

No processo de imputação dos valores omissos poderia ter sido utilizada uma *Random forest imputation*. Este método é uma técnica de *machine learning* que é um procedimento alternativo ao MICE. Este método pode acomodar não linearidades nos dados e interações e não necessita que seja especificado um modelo de regressão.

Para o modelo utilizado, poderiam ser adicionadas componentes não lineares, de forma a enriquecer a explicação do rácio de crédito vencido. Poderiam também ser introduzidas outras variáveis que podem ajudar a explicar o crédito vencido, como por exemplo o PIB ou uma variável ligada ao desemprego.

Por último, e uma vez que foi implementado em Julho de 2018 a Medida do Banco de Portugal seria interessante repetir o estudo com os dados de 2018 e 2019 de forma a verificar se houve uma alteração aos critérios de concessão de crédito. Seria também interessante repetir o estudo quando existisse um período maior de dados após a implementação da Medida de forma a verificar se os impactos dos indicadores macroprudenciais no rácio de crédito vencido se mantêm ou se foram sendo atenuados ao longo do tempo.

6. BIBLIOGRAFIA

- ALMEIDA, Heitor, CAMPELLO, Murillo e LIU, Crocker, (2005). *“The Financial Accelerator Evidence from International Housing Markets”*, 1-30.
- ARREGUI, Nicolas, BENEŠ, Jaromír, KRZNAR, Ivo, MITRA, Srobona e SANTOS, Andre Oliveira (2013). *“Evaluating the Net Benefits of Macroprudential Policy A Cookbook. IMF working paper”*, 13/167.
- Banco de Portugal (2014). *“A política macroprudencial em Portugal objetivos e instrumentos”*.
- Banco de Portugal (2017). *Relatório de Estabilidade Financeira*.
- Banco de Portugal (2018). *“Recomendação do Banco de Portugal no âmbito dos novos contratos de crédito celebrados com consumidores”*.
- Banco de Portugal (2018). *Relatório de Estabilidade Financeira*, 46-53.
- Banco de Portugal (2019). *Acompanhamento da Recomendação macroprudencial sobre novos créditos a consumidores*.
- Basel Committee on Banking Supervision (2010a), *“Guidance for national authorities operating the countercyclical capital buffer”*.
- Basel Committee on Banking Supervision (2006b). *“International Convergence of Capital Measurement and Capital Standards”*.
- Basel Committee on Banking Supervision (2010b). *“An Assessment of the Long-Term Impact of Stronger Capital and Liquidity Requirements”*.
- Basel Committee on Banking Supervision (2017). *“Basel III Finalising post-crisis reforms”*.
- BASTEN, C. e KOCH, C. (2015). *“Higher bank capital requirements and mortgage pricing evidence from the Countercyclical Capital Buffer (CCB)”*, BIS Working Papers 511, Bank for International Settlements.
- Bank for International Settlements (2017). *“BIS Working Papers Macroprudential frameworks, implementation and relationship with other policies, Monetary and Economic Department”*, 94, 374.
- Banco Central Europeu (2015). *Financial Stability Review*, 135-140. Disponível em: <https://www.ecb.europa.eu/pub/pdf/fsr/financialstabilityreview200906en.pdf>.
- BERNANKE, Ben (2010). *“Annual Meeting of the American Economic Association Monetary policy and the housing bubble”*

BELOTTI, Federico, DEB, Partha e MANNING, Willard (2015). “*twopm: Two-part models*” *The Stata Journal* 15, no. 15, 3-20.

BIS, disponível em <https://www.bis.org/publ/bcbs128b.pdf>

BLANCO, Roberto e JIMÉNEZ, Noelia (2018). “*Credit allocation along the business cycle evidence from the latest boom bust credit cycle in Spain*”

BROOKE, Martin, BUSH, Oliver, EDWARDS, Robert, ELLIS, Jas, FRANCIS, Bill, HARIMOHAN, Rashmi, NEISS, Katharine e SIEGERT, Caspar (2015). “*Measuring the macroeconomic costs and benefits of higher UK bank capital requirements, Financial Stability*”, 35, 5-25.

CAMERON, A. Colin e TRIVEDI, Pravin K. (2005), “*Microeconometrics: Methods and Applications*,” Cambridge University Press, New York.

CERUTTI, Eugenio, CLAESSENS, Stijn, e LAEVEN, Luc (2017). “*The Use and Effectiveness of Macroprudential Policies New Evidence, Journal of Financial Stability*”, 28, 203-224.

CRAGG, J. G. (1971), “*Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods*”, *Econometrica*, 39, 829–844.

CROWE, Christopher, DELL’ARICCIA, Giovanni, IGAN, Deniz e RABANAL, Pau (2011). “*How to Deal with Real Estate Booms Lessons from Country Experiences*”. *IMF working paper 11/91*.

Copenhagen Economics –NÆSS-SCHMIDT, Sigurd, JENSEN, Jonas Bjarke, HEEBØLL, Christian e SØRENSEN, Palle (2017): *The role of macroprudential policy in Sweden*.

DEB, Partha e NORTON, Edward C. (2018). “*Modeling Health Care Expenditures and Use*” *Annual Review of Public Health*. 490-503.

DONG-KU, Sung e KIM Doo-Sub (2011) “*A standardization technique to reduce the problem of multicollinearity in polynomial regression analysis*”.

DUAN, N. (1983). “*Smearing estimate A nonparametric retransformation method*”. *Journal of the American Statistical Association* no.78, 605–610.

European Systemic Risk Board (2019). “*A Review of Macroprudential Policy in the EU in 2018*”.

European Systemic Risk Board (2019). “*Macroprudential approaches to non-performing loans*”.

European Systemic Risk Board (2019), “*Vulnerabilities in the residential real estate sectors of the EEA countries*”.

FICHMAN, M. e CUMMINGS, J.N., (2003). “*Multiple imputation for missing data Making the most of what you know*”.

- FMI (2004), *“Compilation guide on financial soundness indicators”*, disponível em: <http://www.imf.org/external/np/sta/fsi/eng/2004/guide/index.htm>.
- FMI (2013), *“Key aspects of macroprudential policy – background paper”*.
- FMI (2013), *“The interaction of monetary policy and macroprudential policies – background paper”*.
- FMI (2014). *“Staff guidance note on macroprudential policy—detailed guidance on instruments”*
- FMI(2017). *“Kingdom of the Netherlands-Netherlands Financial Sector Assessment Program Technical Note-Macroprudential Policy Framework”*.
- FMI(2017). *“Sweden: financial sector assessment program. Technical note—systemic risk oversight framework and management”*. IMF country report no. 17/311.
- FMI(2018). *“Calibration of a debt-service-to-income limit in Romania Evidence from microdata”*. IMF country report no. 18/161
- GROSS, Marco e POBLACIÓN, Javier (2017). *“Assesing the efficacy of borrower-based macroprudential policy using an integrated micro-macro model for European households, European Central Bank”*, 61, pages.
- JIN, Y., LENAIN, P. e O'BRIEN, P. (2014), *“Macroprudential Policy Tools in Norway: Strengthening Financial System Resilience”*, OECD Economics Department Working Papers, No. 1126, OECD
- KAUFMAN, Leonard e ROUSSEEUW, Peter J. (1986). *“Clustering Large Data Sets”*.
- KUTTNER, Kenneth N. e SHIM, Ilhyock (2013). *“Can non-interest rate policies stabilise housing markets? Evidence from a panel of 57 economies”*. Bank for international Settlements Working Papers, 433.
- KZGAR, Ivo e MORSINK, James (2014). *“With Great Power Comes Great Responsibility Macroprudential Tools at Work in Canada”*. IMF working papers, 14/83.
- MALLOWS, Colin Lingwood, *“Some comments on Cp,”* Technometrics, vol. 15, no. 1, November 1973, pp. 661-675
- Marek Zeman (2018), disponível em https://www.cnb.cz/en/public/media_service/press_releases_cnb/2018/20180612_zfs.html
- Monetary Authority of Singapore, 2011, *“Financial stability review”*, Macroeconomic Surveillance Department.
- MONTALVO, Jose G., e RAYA, Josep (2017). *“Constraints on LTV as a macroprudential tool a precautionary tale”*.

- Morris, T.P., White, I.R. e Royston, P. (2014). "Tuning multiple imputation by predictive mean matching and local residual draws". *BMC medical research methodology*, 14(1), p.75.
- NEAGU, Florian, TATARICI, Luminita, e MIHAI, Irina (2015). "Implementing Loan-To-Value and Debt Service-To-Income measures A decade of Romanian experience".
- NEVES, Pedro D. (2010), "As lições da crise e os desafios para a supervisão".
- NEWMAN, D.A., (2014). "Missing data Five practical guidelines. *Organizational Research Methods*", 17(4), pp.372–411.
- RODRIGUES, Raulo e LORENÇO, Rita (2015). "House prices: bubbles, exuberance or something else? Evidence from euro area countries." *Working papers* 2015, Banco de Portugal.
- RUBIN, Donald B. (1987). "Multiple Imputation for Nonresponse in Surveys". John Wiley e Sons, New York.
- PEREIRA, Cristina, RAMALHO, Joaquim J. S., e SILVA, Jacinto Vidigal (2015). "Determinantes do Crédito Vencido nos Bancos de Capital Aberto da OCDE", 50, 21.
- Quivy, R., Campenhoudt, L. V. (2008). "Manual de Investigação em Ciências Sociais Gradiva".
- SCHAFER, J.L (1997). "Analysis of Incomplete Multivariate Data".
- SE, Oh Hwa (2013). "Loan-to-Value as Macro-Prudential Policy Tool Experiences and Lessons of Asian Emerging Countries", *DSF Policy*, 33, 1-12.
- STRUYF, Anja, HUBERT, Mia, ROUSSEEUW, Peter, J. (1970). "Clustering in an Object-Oriented Environment" *Journal of Statistical Software*.
- VAN BUUREN, Stef, BOSHUIZEN HC, KNOOK DL (1999). "Multiple Imputation of Missing Blood Pressure Covariates in Survival Analysis." *Statistics in Medicine*, 18, 681-694.
- VAN BUUREN, Stef, BOSHUIZEN HC, KNOOK DL (1999). "Multiple Imputation of Missing Blood Pressure Covariates in Survival Analysis". *Statistics in Medicine*, 18, 681-694.
- VAN BUUREN, Stef e OUDSHOORN, C.G.M. (2000). "Multivariate Imputation by Chained Equations. MICE V1.0 User's Manual."
- VAN BUUREN, Stef e GROOTHUIS-OUUDSHOORN, Karin (2011). "mice: Multivariate Imputation by Chained Equations in R". *Journal of Statistical Software*, volume 45.
- Wulff J and Ejlskov L, "Multiple Imputation by Chained Equations in Praxis Guidelines and Review" *The Electronic Journal of Business Research Methods* Volume 15 Issue 1 2017, (pp41-56).
Disponível em: www.ejbrm.com

7. ANEXOS

Anexo I - Glossário de variáveis

Anexo II - Métodos de cálculo

Anexo III - Imputação multivariada através de equações em cadeia (Package MICE)

Anexo IV - Clustering Large Applications (CLARA)

Anexo V - Estatísticas descritivas

Anexo VI - Matriz de correlações

Anexo VII - Número ótimo de *clusters* utilizando o método da silhueta média

Anexo VIII - Regressão de duas partes

7.1. ANEXO I - GLOSSÁRIO DE VARIÁVEIS

Nome da variável	Descrição da variável
<i>AnoCelebração</i>	Ano em que o contrato de crédito foi celebrado, ou seja, assinado pelas partes.
<i>MntInicial</i>	Corresponde ao montante total do crédito definido como o limite máximo ou total dos montantes disponibilizados nos termos do contrato de crédito.
<i>MntDivida</i>	Montante do capital em dívida em situação regular no final do período de referência. No caso de contratos em incumprimento, não devem ser incluídas as prestações vencidas e não pagas, nem os encargos decorrentes do incumprimento.
<i>RendAnualBrutoDataIni</i>	Valor do rendimento anual bruto do(s) mutuário(s) na data da celebração do contrato.
<i>PrecoAqHipotecado</i>	Valor da transação do bem hipotecado (e.g. valor da escritura).
<i>ValorPrimeiraAval</i>	Valor do bem hipotecado de acordo com primeira avaliação feita por avaliador independente, externo ou interno.
<i>TAN</i>	Valor da taxa anual nominal aplicável no final do período de referência.
<i>Incumprimento_2017</i>	Se o contrato se encontrar em situação de incumprimento no final do período de referência (2017) preenchido com "1", caso contrário preenchido com "0".
<i>Maturidade</i>	Diferença, em anos, entre a data de celebração do contrato e a data de termo do contrato.
<i>ValorPrimeiraPrest</i>	Valor da prestação mensal na data da celebração do contrato incluindo capital, juros, comissões e despesas (não incluindo seguros).
<i>O_SII</i>	Se o contrato for celebrado com uma das instituições consideradas de importância sistémica preenchido com "1", caso contrário preenchido com "0"
<i>DifAnocelAnoref</i>	Diferença entre o ano de celebração do contrato e o ano em que os dados foram reportados, em anos.
<i>LTlo</i>	Rácio entre o montante do empréstimo e o rendimento do mutuário.
<i>DSTI</i>	Rácio entre o valor da primeira prestação paga pelo mutuário e o seu rendimento.
<i>LTV</i>	Máximo entre o rácio entre o montante do empréstimo e o valor do ativo dado em colateral e o rácio entre o montante do empréstimo e o preço de aquisição do ativo dado em colateral.
<i>RCV</i>	Rácio entre o montante em dívida e o montante inicial para os contratos de crédito, que estejam em 2017, em incumprimento.

7.2. ANEXO II - MÉTODOS DE CÁLCULO

Loan-To-Value na origem com base na primeira avaliação (LTVoa)

$$\frac{\text{Montante Inicial}}{\sum \text{Valor primeira avaliação}_i} \times 100$$

Loan-To-Value na origem com base no preço de aquisição (LTVop)

$$\frac{\text{Montante Inicial}}{\sum \text{Preço de Aquisição Hipotecado}_i} \times 100$$

Em que i corresponde a todos os empréstimos que sejam colateralizados pela mesma garantia hipotecária.

Estes dois indicadores fornecem informação sobre as potenciais perdas para o sistema bancário em caso de incumprimento do crédito, monitorizando o comportamento dos bancos face ao risco no momento de concessão de crédito.

LTV

$$\max(LTVoa; LTVop)$$

Este indicador servirá e *proxy* ao LTV utilizado na Recomendação do Banco de Portugal.

Loan-To-Income na origem (LTIo)

$$\frac{\text{Montante Inicial}}{\text{Rendimento anual bruto data Inicial}} \times 100$$

Debt Service-To-Income na origem (DSTI)

$$\left(\frac{\text{Valor primeira prestação}}{\text{Rendimento anual bruto data início}} \right) \times 100$$

Estes dois indicadores fornecem indicações sobre a probabilidade de incumprimento do crédito no momento da concessão de crédito.

Rácio de crédito vencido

$$\frac{\text{Montante em dívida do crédito em incumprimento}}{\text{Total de Crédito Concedido}}$$

Este indicador apenas é calculado se o contrato se encontrar em incumprimento.

7.3. ANEXO III - IMPUTAÇÃO MULTIVARIADA ATRAVÉS DE EQUAÇÕES EM CADEIA (PACKAGE MICE)

A função utilizada para a imputação multivariada, em R, é a seguinte

```
mice(data, m = 5, method = NULL, maxit = 5)
```

Onde

Data Matriz de dados que contém informação incompleta.

m Número de imputações. Se não for preenchido *m* toma o valor 5.

Method Determina o método de imputação a ser utilizado. Pode ser definido um método igual para todas as variáveis que necessitam de imputação ou o mesmo método para todas.

Maxit Indica o número de iterações a serem efetuadas. Se não for definido o valor *default* é 5.

Avaliação da qualidade das imputações

7.4. ANEXO IV - CLUSTERING LARGE APPLICATIONS (CLARA)

A função utilizada para a análise de *clusters* CLARA, em R, é a seguinte

$$\text{clara}(x, k, \text{metric} = \text{"euclidean"}, \text{stand} = F, \text{samples} = a, \text{sampsiz} = 40+2*k).$$

Onde

x Base de dados de *input*.

k Número de *clusters* a serem considerados.

metric Quando todas as variáveis são numéricas a métrica especifica se as distâncias são computadas através de distância euclidiana ou da distância de Manhattan.

stand As variáveis são normalizadas a unidade zero com a opção *stand=T*. Esta opção é somente utilizada quando todas as variáveis são numéricas. A normalização é útil para evitar que as escolhas dependam das unidades das variáveis.

samples Número de amostras derivados da base de dados original.

sampsiz Número de objetos em cada amostra. Se não for definido o *default* é de $40+2*k$.

7.5. ANEXO V - ESTATÍSTICAS DESCRITIVAS

Tabela 10 - Estatísticas Descritivas para as variáveis relativas ao período de 2005 a 2007

	<i>Média</i>	<i>Mediana</i>	<i>Desvio-Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>
<i>AnoCelebracao</i>	2 006,27	2 006,00	0,76	2 005,00	2 007,00
<i>MntInicial</i>	91 849,03	85 000,00	56 846,96	1 113,50	1 167 187,00
<i>MntDivida</i>	66 426,01	62 841,53	42 793,32	76,22	827 831,71
<i>RendAnualBrutoDataIni</i>	28 577,76	21 774,20	26 869,33	5 303,00	979 286,76
<i>PrecoAqHipotecado</i>	108 167,51	96 800,00	71 913,92	5 100,00	1 800 000,00
<i>ValorPrimeiraAval</i>	161 973,85	139 600,00	90 262,70	5 250,00	1 850 000,00
<i>TAN</i>	0,54	0,27	0,79	0,00	11,00
<i>Incumprimento_2017</i>	0,04	0,00	0,19	0,00	1,00
<i>Maturidade</i>	35,41	38,00	7,82	10,00	56,00
<i>ValorPrimeiraPrest</i>	431,55	392,98	267,56	27,97	1 592,37
<i>O_SII</i>	0,52	1,00	0,50	0,00	1,00
<i>DifAnocelAnoref</i>	10,73	11,00	0,76	10,00	12,00
<i>LTio</i>	5,42	4,64	4,11	0,02	173,35
<i>DSTI</i>	30,14	25,54	22,47	0,11	857,92
<i>LTV</i>	91,55	98,89	28,00	3,09	300,00
<i>RCV</i>	0,03	0,00	0,15	0,00	0,99

Tabela 11 - Estatísticas Descritivas para as variáveis relativas ao período de 2008 a 2014

	<i>Média</i>	<i>Mediana</i>	<i>Desvio-Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>
<i>AnoCelebracao</i>	2 009,23	2 009,00	1,61	2 008,00	2 014,00
<i>MntInicial</i>	95 265,28	85 286,25	61 604,56	1 500,00	1 867 863,30
<i>MntDivida</i>	75 655,01	69 163,36	50 109,78	44,49	1 146 355,70
<i>RendAnualBrutoDataIni</i>	32 690,44	24 897,37	30 959,46	6 000,00	941 812,80
<i>PrecoAqHipotecado</i>	114 211,64	100 000,00	79 799,60	5 037,50	2 000 000,00
<i>ValorPrimeiraAval</i>	166 661,90	143 000,00	101 950,58	5 250,00	3 550 000,00
<i>TAN</i>	0,96	0,57	1,13	0,00	8,73
<i>Incumprimento_2017</i>	0,03	0,00	0,16	0,00	1,00
<i>Maturidade</i>	35,29	38,00	8,73	5,00	51,00
<i>ValorPrimeiraPrest</i>	424,75	374,89	274,74	24,22	1 592,41
<i>O_SII</i>	0,53	1,00	0,50	0,00	1,00
<i>DifAnocelAnoref</i>	7,77	8,00	1,61	3,00	9,00
<i>LTio</i>	5,13	4,06	66,10	0,03	11 208,72
<i>DSTI</i>	27,30	20,38	249,15	0,13	42 161,51
<i>LTV</i>	93,68	99,35	34,70	1,42	300,00
<i>RCV</i>	0,02	0,00	0,14	0,00	1,00

Tabela 12 - Estatísticas Descritivas para as variáveis relativas ao período de 2015 a 2017

	<i>Média</i>	<i>Mediana</i>	<i>Desvio-Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>
<i>AnoCelebracao</i>	2 016,26	2 016,00	0,79	2 015,00	2 017,00
<i>MntInicial</i>	108 378,66	92 800,00	78 262,18	570,00	2 649 000,00
<i>MntDivida</i>	103 158,55	89 003,93	75 344,68	0,00	2 588 149,70
<i>RendAnualBrutoDataIni</i>	35 226,41	23 208,59	42 637,11	5 005,96	955 693,00
<i>PrecoAqHipotecado</i>	140 508,44	117 000,00	110 973,05	6 000,00	3 600 000,00
<i>ValorPrimeiraAval</i>	169 753,98	140 000,00	127 527,48	5 200,00	3 200 000,00
<i>TAN</i>	1,86	1,55	1,05	0,00	8,50
<i>Incumprimento_2017</i>	0,00	0,00	0,05	0,00	1,00
<i>Maturidade</i>	35,88	37,00	8,69	3,00	51,00
<i>ValorPrimeiraPrest</i>	391,52	344,12	240,67	14,03	1 593,68
<i>O_SII</i>	1,00	1,00	0,05	0,00	1,00
<i>Dif_Ano_cel_Ano_ref</i>	0,74	1,00	0,79	0,00	2,00
<i>LTlo</i>	4,69	4,05	3,40	0,01	67,45
<i>DSTI</i>	20,68	17,73	15,37	0,17	280,41
<i>LTV</i>	84,03	89,29	22,83	0,40	298,84
<i>RCV</i>	0,00	0,00	0,05	0,00	1,00

7.6. ANEXO VI - MATRIZ DE CORRELAÇÕES

Tabela 13 - Matriz de correlações para o período de 2005 a 2007

	AnoCelebracao	MntInicial	MntDivida	RendAnualBrutoDataIni	PrecoAqHipotecado	ValorPrimeiraAval	TAN	Incumprimento_2017	Maturidade	ValorPrimeiraPrest	O_SII	DifAnocelAnoref	LTlo	DSTlo	MAXLTV	RCV
AnoCelebracao	1,00															
MntInicial	-0,04	1,00														
MntDivida	0,00	0,91	1,00													
RendAnualBrutoDataIni	0,11	0,30	0,19	1,00												
PrecoAqHipotecado	-0,06	0,83	0,72	0,32	1,00											
ValorPrimeiraAval	0,05	0,59	0,47	0,42	0,68	1,00										
TAN	-0,19	-0,10	-0,06	-0,14	-0,08	-0,19	1,00									
Incumprimento_2017	-0,01	-0,01	0,02	-0,01	-0,02	-0,03	0,19	1,00								
Maturidade	0,00	0,11	0,33	-0,20	0,02	-0,11	0,00	0,03	1,00							
ValorPrimeiraPrest	0,04	0,78	0,71	0,24	0,64	0,39	-0,08	-0,01	0,02	1,00						
O_SII	-0,36	0,19	0,19	-0,18	0,21	-0,07	0,21	0,00	0,15	0,25	1,00					
DifAnocelAnoref	-1,00	0,04	0,00	-0,11	0,06	-0,05	0,19	0,01	0,00	-0,04	0,36	1,00				
LTlo	-0,18	0,21	0,26	-0,40	0,17	0,05	0,09	0,02	0,28	0,12	0,27	0,18	1,00			
DSTlo	-0,08	0,15	0,19	-0,41	0,10	-0,04	0,10	0,01	0,21	0,30	0,31	0,08	0,86	1,00		
MAXLTV	0,03	-0,01	-0,01	0,00	-0,01	0,00	0,01	0,01	-0,01	-0,02	-0,04	-0,03	-0,01	-0,02	1,00	
RCV	-0,01	-0,01	0,03	-0,02	-0,02	-0,03	0,19	0,99	0,06	-0,01	0,00	0,01	0,03	0,02	0,01	1,00

Tabela 14 - Matriz de correlações para o período de 2008 a 2014

	AnoCelebracao	MntInicial	MntDivida	RendAnualBrutoDataIni	PrecoAqHipotecado	ValorPrimeiraAval	TAN	Incumprimento_2017	Maturidade	ValorPrimeiraPrest	O_SII	DifAnocelAnoref	LTlo	DSTlo	MAXLTV	RCV
AnoCelebracao	1,00															
MntInicial	0,03	1,00														
MntDivida	0,10	0,94	1,00													
RendAnualBrutoDataIni	-0,03	0,29	0,21	1,00												
PrecoAqHipotecado	0,07	0,83	0,76	0,33	1,00											
ValorPrimeiraAval	-0,05	0,62	0,53	0,39	0,72	1,00										
TAN	0,57	-0,08	-0,02	-0,09	-0,05	-0,17	1,00									
Incumprimento_2017	-0,05	-0,02	0,00	0,00	-0,03	-0,02	0,09	1,00								
Maturidade	0,03	0,13	0,29	-0,20	0,02	-0,10	0,00	0,03	1,00							
ValorPrimeiraPrest	-0,02	0,68	0,63	0,19	0,56	0,33	0,00	0,00	0,01	1,00						
O_SII	0,48	0,13	0,18	-0,14	0,16	-0,06	0,37	-0,02	0,17	0,23	1,00					
DifAnocelAnoref	-1,00	-0,03	-0,10	0,03	-0,07	0,05	-0,57	0,05	-0,03	0,02	-0,48	1,00				
LTlo	0,00	0,02	0,02	-0,02	0,01	0,02	0,00	0,00	0,02	0,01	0,00	0,00	1,00			
DSTlo	0,00	0,01	0,02	-0,03	0,01	0,01	0,00	0,00	0,02	0,03	0,01	0,00	1,00	1,00		
MAXLTV	-0,18	0,01	0,00	-0,01	-0,07	-0,03	-0,11	0,01	0,05	0,04	-0,06	0,18	0,02	0,02	1,00	
RCV	-0,04	-0,01	0,00	-0,01	-0,03	-0,02	0,10	0,99	0,04	0,00	-0,01	0,04	0,00	0,00	0,01	1,00

Tabela 15 - Matriz de correlações para o período de 2015 a 2017

	AnoCelebracao	MntInicial	MntDivida	RendAnualBrutoDataIni	PrecoAqHipotecado	ValorPrimeiraAval	TAN	Incumprimento_2017	Maturidade	ValorPrimeiraPrest	O_SII	DifAnocelAnoref	LTlo	DSTlo	MAXLTV	RCV
AnoCelebracao	1,00															
MntInicial	0,04	1,00														
MntDivida	0,06	0,96	1,00													
RendAnualBrutoDataIni	0,00	0,44	0,42	1,00												
PrecoAqHipotecado	0,05	0,86	0,82	0,51	1,00											
ValorPrimeiraAval	0,04	0,75	0,71	0,49	0,88	1,00										
TAN	-0,33	-0,21	-0,20	-0,10	-0,21	-0,21	1,00									
Incumprimento_2017	-0,05	-0,02	-0,02	-0,02	-0,02	-0,02	0,10	1,00								
Maturidade	0,03	-0,01	0,02	-0,27	-0,13	-0,14	-0,02	0,01	1,00							
ValorPrimeiraPrest	-0,05	0,69	0,66	0,40	0,61	0,53	-0,14	-0,01	-0,22	1,00						
O_SII	0,00	0,01	0,01	0,01	0,02	0,01	-0,04	-0,01	0,04	0,00	1,00					
DifAnocelAnoref	-1,00	-0,04	-0,06	0,00	-0,05	-0,04	0,33	0,05	-0,03	0,05	0,00	1,00				
LTlo	0,05	0,21	0,20	-0,37	0,09	0,08	-0,02	0,02	0,30	0,11	-0,05	-0,05	1,00			
DSTlo	-0,06	0,06	0,05	-0,37	-0,01	-0,03	0,06	0,03	0,12	0,23	-0,06	0,06	0,83	1,00		
MAXLTV	-0,02	0,07	0,05	-0,14	-0,26	-0,21	0,07	0,02	0,28	0,01	-0,01	0,02	0,22	0,16	1,00	
RCV	-0,05	-0,02	-0,02	-0,02	-0,02	-0,02	0,10	1,00	0,01	-0,01	-0,01	0,05	0,02	0,03	0,02	1,00

7.7. ANEXO VII - NÚMERO ÓTIMO DE CLUSTERS UTILIZANDO O MÉTODO DA SILHUETA MÉDIA

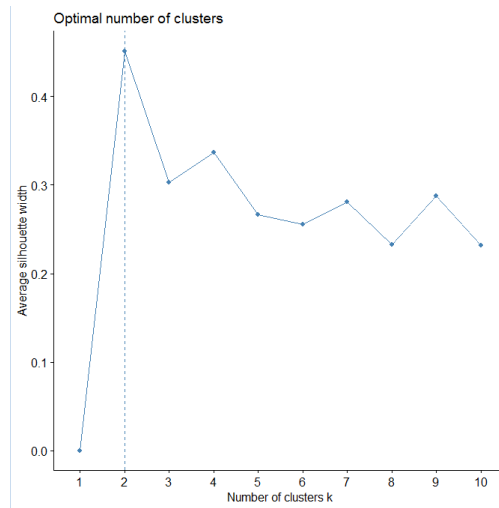


Figura 16 - Número ótimo de *clusters* para o período de 2008 a 2014

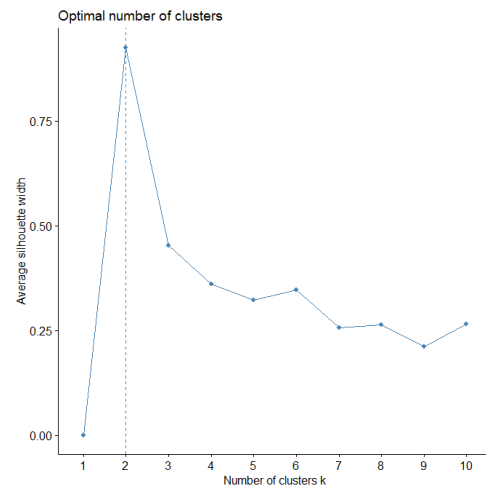


Figura 17 - Número ótimo de *clusters* para o período de 2015 a 2017

7.8. ANEXO VIII - REGRESSÃO DE DUAS PARTES

Sinais esperados

Grupo de instituições em que foi solicitado o crédito

- O-SII (-)

Imóvel dado como colateral

- PreçoAqHipotecado (-)
- LTV (+)

Condições dos mutuários

- DSTI (+)
- RendAnualBrutoDataIni (-)

Condições do contrato no momento da contratação

- Maturidade (+)
- MntInicial (+)
- TAN (+)

Características do contrato à data de reporte do mesmo

- DifAnocelAnoref (+)



Rácio de Crédito Vencido