

UNIVERSIDADE PRESBITERIANA MACKENZIE
Centro de Ciências Sociais Aplicada
Programa de Pós-Graduação em Administração de Empresas

SANDRA ALMEIDA SILVA

**ESTUDO DE RISCO DE CRÉDITO EM OPERAÇÕES DE CARTÃO DE
CRÉDITO USANDO VARIÁVEIS MACROECONÔMICAS E
TÉCNICAS DE ANÁLISE DE SOBREVIVÊNCIA**

São Paulo
2012

Sandra Almeida Silva

Estudo de Risco de crédito em operações de cartão de crédito usando
variáveis macroeconômicas e técnicas de Análise de Sobrevivência

Dissertação apresentada ao Programa de Pós-Graduação em Administração de Empresas da Universidade Presbiteriana Mackenzie, para a obtenção do título de Mestre em Administração de Empresas.

Orientador: Prof. Dr. Herbert Kimura

São Paulo
2012

SANDRA ALMEIDA SILVA

**ESTUDO DE RISCO DE CRÉDITO EM OPERAÇÕES DE CARTÃO DE
CRÉDITO USANDO VARIÁVEIS MACROECONÔMICAS E
TÉCNICAS DE ANÁLISE DE SOBREVIVÊNCIA**

Dissertação apresentada ao Programa de Pós-Graduação em Administração de Empresas da Universidade Presbiteriana Mackenzie, para obtenção do título de Mestre em Administração de Empresas.

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Herbert Kimura – Orientador
Universidade Presbiteriana Mackenzie

Prof. Dr. Leonardo Basso
Universidade Presbiteriana Mackenzie

Prof. Dr. Eduardo Kazuo Kayo
(FEA-RP-USP)

À minha mãe, pela infinita sabedoria.
Aos meus filhos Beatriz e Pedro, razões da minha vida.

AGRADECIMENTOS

A energia criadora do universo pela força e coragem concedida durante todo o percurso desta caminhada.

Ao Professor Dr. Herbert Kimura pela sábia orientação e que apesar das minhas limitações soube conduzir-me firmemente na conclusão deste trabalho. Agradeço também sua admirável paciência e bondade acadêmica.

Aos professores, em especial, ao Professor Dr. Leonardo Basso pelas preciosas discussões, á todos os funcionários da Pós-Graduação Mackenzie, pela colaboração e pronto atendimento à realização deste trabalho.

Às minhas irmãs, minha família e ao Marcos, meu companheiro, pelo incondicional apoio em todas as horas deste caminho.

Aos meus queridos amigos pela compreensão, incentivo e confiança nos momentos mais difíceis deste processo.

Ao Rafael P., colega de curso, pelo companheirismo e amizade.

Ao MACKPESQUISA – Fundo Mackenzie de Pesquisa que financiou, em parte, este trabalho.

A CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela concessão da Bolsa CAPES-PROSUP - Mod. I.

“Se, a princípio, a ideia não é absurda, então não há esperança para ela”.
Albert Einstein.

RESUMO

Nas duas últimas décadas, devido à expansão do mercado de crédito, tornou-se necessário o desenvolvimento de modelos de classificação de risco que não somente capturassem a probabilidade de inadimplência, mas que estimassem “quando” ocorreria o momento de *default*. Este estudo investiga se a probabilidade de default é afetada pelas condições gerais da economia ao longo do tempo, utilizando a técnica de Análise de Sobrevivência. Usando dados de titulares de cartões de crédito de uma financeira ligada a um banco, analisaram-se potenciais influências de variáveis de cadastro e de variáveis macroeconômicas no tempo de sobrevivência, ou seja, de não inadimplência. Os resultados sugerem que modelos que incorporam variáveis macroeconômicas possibilitam melhorar o nível de explicação do tempo de sobrevivência, comparativamente a modelos que utilizam somente dados de cadastro. Considerando variáveis de cadastros, os testes estatísticos apontaram o sexo feminino possui maior sobrevivência, indicando que homens tendem a entrar em inadimplência mais cedo na amostra estudada. Com relação às variáveis macroeconômicas, o nível de desemprego mensal tem maior relevância na explicação de inadimplência.

Palavras-chave: Análise de Sobrevivência. Inadimplência. Variáveis Macroeconômicas.

ABSTRACT

In the last two decades due to expansion of the credit market, it became necessary to develop risk classification models that not only capture the probability of default, but also “when” is the time of default. This study investigated the probability of default is affected by general economic conditions over time due to the use of Survival Analysis technique. Using data from credit card holders of a financial bank the objective was to analyze potential influences of variables of record, register and macroeconomic variables on survival time, it means, not default. The results suggest that models incorporating macroeconomic factors allow the improvement of the explanation of survival time compared to models that utilize only data register/records. Considering records/registers variables, statistical tests showed that females had higher survival, indicating that men tend to go into default earlier in the studied sample. Regarding the macroeconomic variables, the level of monthly unemployment has more relevance in explaining default.

Keywords: Survival Analysis, Default, Macroeconomics Variables.

LISTA DE FIGURAS

Figura 1: Curvas de Kaplan-Meier.....	62
Figura 2: Curvas de Kaplan-Meier: estratificado por fator.....	64
Figura 3: Probabilidade de sobrevivência Log-Log.....	69
Figura 4: Teste de Resíduos de Schoenfeld do modelo de Cox com interação de Limite e Renda Confirmada.....	71

LISTA DE TABELAS

Tabela 1: Cronograma de adequação as regras da Basiléia II.....	22
Tabela 2: Variáveis Macroeconômicas utilizadas na modelagem de risco de crédito para consumo.....	35
Tabela 3: Tabela de sobrevivência.....	43
Tabela 4: Variáveis da amostra.....	49
Tabela 5: Variáveis da pesquisa.....	50
Tabela 6: Variáveis macroeconômicas.....	52
Tabela 7: Matriz dos dados da amostra.....	58
Tabela 8: Análise descritiva dos dados da amostra.....	59
Tabela 9: Tempo de sobrevivência de cada individuo.....	60
Tabela 10: Estimando função de sobrevivência.....	61
Tabela 11: Resultados Teste Logrank.....	65
Tabela 12: Resultados modelo de Regressão Paramétrico.....	67
Tabela 13: Resultados do modelo de risco proporcional de Cox – limite e renda confirmada.....	70
Tabela 14: Teste de Resíduos Schoenfeld.....	73

Tabela 15: Resultados do modelo de risco proporcional de Cox com variáveis macroeconômicas.....	75
---	----

SUMÁRIO

1	INTRODUÇÃO	8
1.1	Relevância da Pesquisa	12
1.2	Objetivos do Projeto	12
2	REFERENCIAL TEÓRICO	14
2.1	Risco de Crédito.....	14
2.2	Regulamentações sobre crédito: Basiléia I e Basiléia II.....	17
2.2.1	Regras do Banco Central do Brasil relacionadas a crédito	20
2.3	Formas de análise em modelos de classificação de risco de crédito.....	22
2.3.1	Modelos de concessão de crédito	23
2.3.2	Modelos de concessão: Principais modelos de carteira de crédito.....	26
2.4	Estudos empíricos sobre análise de crédito para empréstimos ao consumo.....	31
2.4.1	Empréstimos por meio de cartão de crédito.....	32
2.4.2	Variáveis explicativas utilizadas na modelagem de risco de crédito.....	35
2.5	Inadimplência.....	38
3	METODOLOGIA	40
3.1	Problema de Pesquisa	40
3.2	Técnica de Análise	40
3.2.1	Análise de Sobrevivência.....	40
3.2.2	Censura.....	44
3.2.3	Funções do Tempo de Sobrevivência e Funções de risco	45
3.2.4	Estimação da função de sobrevivência	45
3.2.5	Modelo de riscos proporcionais de Cox.....	46
3.2.6	Covariáveis dependentes do tempo	47
3.2.7	Modelo de Cox - Estratificado.....	47
3.3	Base de Dados	48
3.3.1	Variáveis macroeconomicas	51
3.4	Modelo matemático	53
3.4.1	Forma de análise dos resultados.....	56
4	RESULTADOS	57
4.1	Função de sobrevivência.....	59

4.2	Estimativas de sobrevivência de Kaplan-Meier.....	60
4.2.1	Estimativas de sobrevivência de Kaplan-Meier por fatores.....	62
4.3.2	Teste estatístico de <i>Log-Rank</i>	63
4.4.	Modelo de Regressão Paramétrico.....	65
4.5.	Teste de proporcionalidade do Modelo de Cox – Log Log.....	68
4.5.1	Teste Schoenfeld - Análise dos Resíduos	70
4.5.2	Modelo semiparamétrico de risco proporcional de Cox	72
4.5.3	Modelo de riscos proporcionais de Cox com variáveis macroeconômicas.....	74
5 - CONSIDERAÇÕES FINAIS		77
6 - REFERÊNCIAS BIBLIOGRÁFICAS.....		81

1 INTRODUÇÃO

Desde que a sociedade começou a tomar decisões com a ajuda dos números, a modelagem preditiva de risco de crédito tem sido um desafio na área de Finanças.

Fisher (1936) foi um dos principais percussores na área estatística de regressão logística, utilizando em seus modelos a técnica mais conhecida como *Credit Scoring*, metodologia em duas categorias: modelos de concessão (*application scoring*) e modelos comportamentais (*behavioural scoring*) objetivando explorar hipóteses causais do comportamento dos tomadores de crédito em relação ao consumo e, portanto instituições financeiras e agentes reguladores têm utilizado ao longo do tempo tais modelos para avaliar e estimar o risco individual ou possíveis perdas nas carteiras de crédito.

Porém Thomas (2009) relata que apesar do consumo de crédito ser um fenômeno surgido nos últimos 50 anos, e está relacionado à necessidade de financiamento de crédito desde o desenvolvimento do mercado de crédito em 1920, com Henry Ford até os dias atuais. O autor afirma que a forma desordenada de acesso ao crédito tanto pelos agentes econômicos como pelos consumidores enfatiza a necessidade do desenvolvimento de modelos de risco de crédito mais eficientes e modernos adequados a esta crescente demanda por crédito.

Por outro lado, os bons resultados trazidos pela oferta de crédito estimula também uma reação quase que sintomática como resposta, a inadimplência. O Banco Central do Brasil, em seu Relatório de Estabilidade Financeira (Banco Central do Brasil, 2011), relata que para equacionar as necessidades de gestão de risco em modelos eficientes de classificação de risco de crédito, é necessária uma força conjunta entre regulados e reguladores, no intuito de absorver o impacto da Basileia II nas operações de crédito, de maneira ordenada e não restritiva à expansão do crédito.

As instituições financeiras possuem grande importância no desenvolvimento do mercado de crédito por possuírem a função credora e mais importante por representarem a credibilidade e segurança do sistema financeiro. Tal importância é percebida, por exemplo, no mercado brasileiro, onde a participação das instituições privadas nacionais e estrangeiras é de 40,87% e 17,81%, respectivamente, conforme o Banco Central (2011).

De acordo com Caouette et al. (1998, p.104) a prioridade no desenvolvimento de modelos de classificação de risco é primeiramente para atender às instituições financeiras que necessitam de instrumentos de gestão e controle de risco de crédito que colaborem para a eficiência bancária, principalmente nas rotinas de classificação de risco interno, previsão de inadimplência, e concessão de crédito.

Tal necessidade é percebida no mercado de crédito pelo risco de insolvência percebido pelas instituições financeiras caso não possuam instrumentos de avaliação de risco capazes de avaliarem grandes volumes de crédito não honrados e, por conseguinte, não se proteger do risco de inadimplência.

Recentemente diversos pesquisadores (THOMAS et al. 2005; ANDREEVA et al. 2003; BELLOTTI e CROOK, 2007) têm se empenhado no desenvolvimento de modelos de risco de crédito que possam capturar a probabilidade de *default* resultante de carteiras de empréstimos e, conseqüentemente, proporcionar-lhes melhoras na eficiência preditiva dos modelos de risco e rapidez nas decisões de concessão de crédito.

Entretanto tais modelos possuem algumas limitações observadas por Mester (1977) e Hand e Kelly (2001) ao longo dos últimos anos, por exemplo, uma das importantes limitações é considerarem apenas os clientes que foram bons ou maus pagadores em uma determinada amostra e, portanto, não contemplam a previsão de padrão em uma carteira de empréstimos, principalmente para consumo, em que se têm grandes volumes de dados. Limitações também pesquisadas por Hian e Chwee (2006) e Verstraeten e Poel (2005) em que advertem sobre o viés de amostragem para estudos em grandes amostras.

Thomas e Stepanova (2005), Andreeva (2006) e Malik e Thomas (2009) testaram a inclusão de variáveis macroeconômicas em seus modelos de risco de crédito informam que há uma classe de modelos de risco proporcional, Cox (1972), nos quais a técnica estatística de Análise de Sobrevivência é aplicada como uma alternativa flexível na modelagem do risco de crédito por incorporar variáveis como resposta, por exemplo: tempo de sobrevivência de um indivíduo até um determinado evento de interesse, conforme proposto por Kachman (1999).

Neste contexto, este trabalho pretende demonstrar através da aplicação do modelo de classificação de risco de crédito de Bellotti e Crook (2007) se fatores macroeconômicos incluídos em modelo de risco de crédito através da técnica estatística de Análise de Sobrevivência podem contribuir para o desenvolvimento de modelos de risco de crédito e que possam estimar a probabilidade futura de *default* em carteira de crédito ao consumo.

Os dados analisados são de indivíduos que obtêm crédito na forma de empréstimos por meio de cartões de crédito, configurando crédito ao consumo, e o Evento de Interesse é o não pagamento da fatura do cartão de crédito após 60 dias de atraso, sendo considerado Enquadrado, ou seja, inadimplente, norma aplicada segundo contrato do emissor do cartão de crédito e o usuário.

Primeiramente os dados serão estratificados pelo fator Sexo e analisados através da técnica estatística de Análise de Sobrevivência, onde a variável resposta é o tempo até a ocorrência de um determinado evento de interesse, também chamado de Tempo de Sobrevivência, Kleinbaum e Klein (2006); ou Tempo de Falha, Colosimo e Giolo (2006).

O estimador utilizado para a observação dos dados, inclusive os censurados, será o Kaplan Meier, estimador não paramétrico que estimará a função de sobrevivência, considerando todos os indivíduos da amostra. Em seguida o teste de *Logrank* irá comparar o número de eventos observados em cada grupo com o número de eventos esperados ordenando os tempos de "falhas" dos indivíduos nos dois grupos. O Objetivo é testar a significância estatística da relação de igualdade entre as curvas de sobrevivência na amostra estratificada pelo fator Sexo.

Em um segundo momento será a modelagem das variáveis selecionadas da amostra nos modelos escolhidos, para tanto será necessário testar a adequabilidade dos modelos a serem utilizados. O objetivo será verificar se o modelo de risco proporcional Cox é apropriado para a inclusão das variáveis, Limite e Renda Confirmada como variáveis que não mudam ao longo do tempo, estabelecendo, portanto uma dependência do tempo.

Para Kleinbaum e Klein (2006) em relação aos modelos probabilísticos, há principalmente dois tipos, os modelos de risco proporcional, por exemplo, o modelo de Cox, Cox (1972) que considera que o risco é constante ao longo do tempo e o modelo Paramétrico onde o pressuposto é que o tempo de sobrevida é acelerado ou desacelerado por um fator constante.

Portanto aplica-se o teste de resíduos Schoenfeld, Kleinbaum e Klein (2006) para verificar se o pressuposto de proporcionalidade é satisfeito. Ou seja, os resíduos não estão correlacionados com o tempo de sobrevivência evidenciando, portanto que o pressuposto de risco proporcional não foi violado.

Após os testes de Schoenfeld o objetivo será analisar as variáveis, Cadastro, primeiramente no modelo de Regressão Paramétrico. A intenção é explorar a relação entre as covariáveis que podem estar relacionadas com o tempo de sobrevivência sob o pressuposto de que o tempo de sobrevida é acelerado ou desacelerado por um fator constante.

As variáveis de Cadastro escolhidas foram: Limite de Crédito e Renda Confirmada, conforme referencial teórico e prática do mercado de crédito para concessão de crédito. A expectativa é que possamos afirmar se tais variáveis são importantes para explicar inadimplência.

Em seguida a amostra será testada no modelo de riscos proporcionais semiparamétrico de Cox. Novamente o objetivo é avaliar se tais covariáveis influenciam o tempo de sobrevivência de cada indivíduo e o efeito dessas variáveis no evento de interesse. Definindo, portanto a função de risco no modelo. A intenção é analisar de acordo com os coeficientes apresentados, e associados à taxa de risco se haverá ou não aumento ou diminuição do risco de inadimplência pelos indivíduos da amostra.

Por último os dados serão modelados no modelo de Cox – Estratificado, ou Estendido Colosimo e Giolo (2006) com a inclusão das variáveis macroeconômicas. A expectativa é demonstrar se a inclusão das variáveis macroeconômicas fornecerá ao modelo resultados estatisticamente significantes em relação à taxa de risco.

A taxa de risco será obtida pela função exponencial do coeficiente das variáveis visando identificar o efeito das mesmas no modelo, e ainda pretende-se observar se as variáveis macroeconômicas consideradas como independentes "x" possam explicar "y" inadimplência.

As variáveis macroeconômicas escolhidas foram: Taxa de Desemprego, Taxa de Juros- Selic e IPCA- (Inflação), a escolha também obedeceu ao critério apresentado no referencial teórico. Tais variáveis serão analisadas em conjunto com as variáveis "Cadastro" (Limite de Crédito e Renda Confirmada), já testadas anteriormente nos modelos paramétrico e semiparamétrico.

A expectativa é que a intensidade de inadimplência dos consumidores é significativamente influenciada por fatores macroeconômicos, conforme estudos apresentados por Malik e Thomas (2009).

1.1 Relevância da Pesquisa

Nos dias atuais devido a novas pesquisas e abordagens na área de risco de crédito surge a necessidade de modelos que sejam capazes de administrar a gestão de risco de crédito das instituições financeiras. A gestão do risco de crédito interno possibilitará às instituições financeiras a administração de suas provisões de dívidas, bem como o tratamento adequado a grandes volumes de dados. Ainda promover o desenvolvimento de modelos de classificação de risco que possam aumentar a capacidade preditiva de *default* dos modelos atuais de *Credit Score*.

A Pesquisa Nacional de Endividamento e Inadimplência do Consumidor (PEIC, 2011) informa que cerca de 58,0% das famílias brasileiras estão endividadas sendo: 21,2% com dívidas em atraso, 7,2 % inadimplentes e 71,8% dos entrevistados afirmaram que o cartão de crédito é um dos principais tipos de dívida.

Desse modo acredita-se que a aplicação do modelo de Bellotti e Crook (2007) em dados brasileiros através da técnica estatística de Análise de Sobrevivência com variáveis macroeconômicas possa ser um exercício produtivo na análise de risco de crédito ao consumo, além é claro de ampliar as discussões acerca do tema.

1.2 Objetivos do Projeto

Através de uma base de dados de operações de crédito de uma instituição financeira serão analisadas funções de sobrevivência por meio de variáveis de Cadastro do indivíduo bem como variáveis macroeconômicas.

O principal objetivo é estudar se as variáveis denominadas Cadastro: Limite de Crédito, Renda Confirmada e Sexo, em conjunto com as variáveis macroeconômicas: IPCA, Taxa de Desemprego e Taxa de Juros - Selic possuem significância estatística para explicar inadimplência, como possíveis variáveis respostas, a partir de técnicas de Análise de Sobrevivência.

É objetivo também desta pesquisa analisar se devido à inclusão de variáveis macroeconômicas no modelo de risco de crédito este fornecerá capacidade preditiva sobre o aumento ou diminuição do risco de inadimplência, e, portanto, sugerindo que alterações nas condições gerais da economia possam aumentar a probabilidade de default dos tomadores de crédito.

2 REFERENCIAL TEÓRICO

Neste capítulo é apresentada a fundamentação teórica sobre risco de crédito. São analisados os principais modelos de classificação de risco de crédito, regulamentações, formas de análises, estudos empíricos e as variáveis explicativas utilizadas na modelagem de risco de crédito para construção de modelos previsores de *default*.

2.1 Risco de crédito

A classificação de crédito surgiu por volta de 1950, com os trabalhos de Fisher (1936), segundo Thomas (2009) em métodos estatísticos de classificação de risco de crédito, sendo que a premissa básica era classificar dados de empréstimos em bons e maus tomadores com o objetivo em fornecer suporte às decisões em relação à concessão de crédito.

O BCBS (1988) - Comitê de Basileia sobre Supervisão Bancária - define que risco de crédito é a possibilidade de um tomador de crédito ou contraparte em falhar no pagamento de suas obrigações em relação aos termos contratados.

O que para Jacobs (2010) risco de crédito seria a perda no potencial creditício do tomador e das contrapartes, o que naturalmente contribui de maneira decisiva na perda do capital econômico das instituições e, portanto, deflagra a possibilidade de falências bancárias, desencadeando o risco sistemático.

A literatura especializada sobre risco de crédito de um modo geral tem sido abordada por premissas pioneiras sobre teorias de finanças, a mais conceituada no mercado financeiro surgiu a partir do artigo *Portfolio Selection*, Markowitz (1952) que analisa o risco a partir do retorno esperado de cada ativo em relação ao mercado. Desse modo observa-se que a racionalidade esperada dos agentes econômicos seria responsável pela construção do paradigma central das teorias tradicionais de finanças, sendo a moderna Teoria das Carteiras de Markowitz um dos trabalhos pioneiros sobre o conceito de risco, aplicado à carteira de crédito.

Entretanto o conceito ‘risco’ nos dias atuais é reforçado devido as recentes crises financeiras onde os agentes econômicos reforçam a busca desta racionalidade em modelos de classificação de risco como, por exemplo, mencionado em Banasik *et al.* (1999), ao definirem classificação de risco de crédito como sistemas construídos para sustentar hipóteses em relação à probabilidade futura de *default* de um tomador de crédito.

Para Thomas *et al.* (2002) tal racionalidade pode-se encontrar também em modelos estatísticos que viabilizem a classificação do risco e a tomada de decisão frente à concessão ou não de empréstimos. Desse modo ampliam-se as discussões acerca de modelos de classificação que possam mensurar o risco no mercado de crédito ao consumo, principalmente em relação ao processo decisório de concessão de crédito conforme visto em Andreeva *et al.*, 2003 e também nos modelos que desenham variáveis comportamentais previsoras de padrões futuros de comportamento, proposto por Baesens *et al.* (2005), ou ainda, o modelo analisado por Malik e Thomas (2005), que relacionam os aspectos macroeconômicos como variáveis no tempo.

Thomas *et al.* (2005) mencionam que os aspectos macroeconômicos de um país possuem grande importância no desenvolvimento de modelos de gerenciamento de risco de crédito principalmente no atendimento aos requerimentos regulatórios e, como resultado da crescente demanda de crédito, sobretudo, na proteção ao sistema financeiro diante das recentes crises financeiras.

Entretanto para Caouette *et al.* (1998) apesar da importância afirmam existir grandes dificuldades no desenvolvimento de modelos de risco de crédito que possam absorver todas as necessidades demandadas por um mercado de crédito em desenvolvimento e expansão, e estas dificuldades se devem principalmente aos novos entrantes no mercado, à diminuição da margem de lucro das instituições financeiras e, por último, ao surgimento de novas teorias de finanças diante da compreensão dos fatores de risco em relação ao mercado de crédito.

Os autores argumentam que as dificuldades e complicações encontradas no desenvolvimento de modelos é fruto, em parte, da grandiosidade do mercado de crédito ao consumidor, não somente em termos monetários, mas no grande número de consumidores envolvidos no processo.

Desse modo a racionalidade esperada dos agentes credores em relação à modelagem de risco de crédito é testada nos últimos anos em relação ao correto entendimento regulatório pelas instituições, (THOMAS *et al.*, 2005), pois a seu ver o excesso de regulamentação pode engessar o desenvolvimento de novos modelos de classificação de risco.

Por outro lado a completa desregulamentação pode incorrer em risco, principalmente o risco sistemático, conforme demonstrado nos estudos macroeconômicos de Tang *et al.* (2007).

Igualmente importantes são os modelos capazes de prever o risco sistemático, capazes de quantificar o capital econômico de uma instituição (JACOBS, 2010) como, por exemplo, os modelos de Vasicek (2002), adotados pelo Novo Acordo da Basileia BSBS (2004). Tais modelos apresentaram uma contribuição técnica valiosa aos modelos de risco de crédito por demonstrarem que a probabilidade de distribuição de perdas, ou seja, a perda esperada (PD) é tema relevante em modelos de classificação de risco em previsões de padrões futuros de comportamento para carteiras de crédito de grandes amostras, no caso, empréstimos ao consumo.

Porém, Caouette *et al.* (1998) mencionam que diante das dificuldades expostas em modelagem de risco de crédito, o volume movimentado no mercado de crédito é fator preponderante no desenvolvimento de modelos de risco de crédito.

Porém, segundo os autores, apesar das dificuldades apresentadas por esta característica, pode-se considerar uma variável previsor de ciclos de negócios na economia de um país, e conseqüentemente traduzir-se em risco aos agentes credores.

Para Bellotti e Crook (2007; 2008) a preocupação é principalmente em relação à utilização de cartões de crédito, como sendo um das principais fontes de recursos para o consumo, os autores informam que a inclusão de variáveis macroeconômicas nos modelos de risco de crédito pode deflagrar uma reação contínua e em cadeia em relação à previsão de inadimplência afetando de maneira generalizada a econômica global.

A título de exemplo, Capozzi (2011), em artigo publicado no jornal *The Wall Street Journal*, e de acordo com o *Federal Reserve*, o crédito ao consumidor norte-americano é de

aproximadamente \$2.14 trilhões, sendo a média norte-americana de dívida familiar de \$75.600, contrapondo a renda média de \$49.800.

Portanto demonstrando que, apesar da economia mundial depender do consumo de bens e serviços, os números indicam que o consumo ocorre com dinheiro emprestado e, principalmente, dinheiro que não pode ser reembolsado, caso haja necessidade de reembolso na íntegra pelos credores.

No Brasil não é diferente, o volume de crédito destinado às famílias brasileiras atingiu R\$737 bilhões, conforme o Banco Central do Brasil (2011), e representa um crescimento semestral de 12%, sendo os financiamentos habitacionais com 24,6% e o crédito consignado, 13%. Este cenário evidencia que o endividamento assume grande importância no comprometimento da renda média dos brasileiros, considerando os 5,1 milhões de novos entrantes no sistema financeiro em 2011, tendo as exposições individuais um aumento de 23% em 12 meses, demonstrando conseqüentemente, o aumento gradativo da exposição ao risco do sistema bancário às pessoas físicas.

Acompanhar a trajetória do endividamento e do comprometimento de renda dos consumidores é crucial não apenas para mensurar o risco de crédito, mas também para verificar o potencial de consumo das famílias e o seu impacto no crescimento econômico (Bellotti e Crook, 2008) e, portanto, investigar os possíveis fatores que possam indicar probabilidades de *default*.

Desse modo, confirma-se o entendimento, mesmo que intuitivo, que à medida que se expande a oferta de crédito, corre-se o risco de afetar fatores macroeconômicos gerais e, portanto contribuir para a inadimplência, principalmente se ocorrer de maneira desordenada (CAOINETTE *et al.*,1998).

2.2 Regulamentações sobre crédito: Basiléia I e Basiléia II

Para Crouhy *et al.* (2000), a expansão do mercado de crédito acarreta o estabelecimento de regulamentação internacional, justificando a necessidade de supervisão bancária, como o Acordo da Basiléia.

O *Bank for International Settlements* (BIS), organização internacional que fomenta a cooperação financeira e monetária internacional, ou seja, o banco para os bancos centrais tem como principal obrigação e objetivo estabelecer esta regulamentação. Por exemplo, estabelecer o correto índice de capital regulamentar para as instituições financeiras, fator de grande importância nos dias atuais para as instituições financeiras. Tais obrigações entre outras tem sido realizadas desde 1988 até as mais recentes revisões como o Novo Acordo (BCBS, 2003) e posterior implementação em 2007.

O Relatório da Basileia (2004) estabelece que o percentual mínimo para alocação do capital é fundamental para o gerenciamento de risco de crédito e que adequar à exigência de capital regulatório é uma consequência à tendência de internacionalização das atividades bancárias e propicia proteção ao sistema financeiro em um mundo globalizado.

O Comitê endossou tais afirmações na tentativa de implementar políticas econômicas, com o objetivo de controlar a disponibilidade de crédito ao consumo no país, conforme impacto previsto por Jackson (1999).

Porém, segundo o Acordo BIS (2006), *Bank for International Settlements*, a intenção é incentivar a adoção do capital regulamentar para as instituições financeiras em duas abordagens, sendo elas: IRB básica (*Internal Rating Based*) e IRB Avançada, (*Internal Rating Based Advanced*) com padronização por meio das agências de *rating*, com o objetivo principal de mensurar a probabilidade de *default*, Vasicek (2002); Exposição ao *Default* (EAD), Schuermann T. (2004); *Loss Given Default* (LGD), Jacobs e Karagozoglou (2010) e o *Maturity* (M), He. Z e Xiong W. (2010) em relação à perda esperada e inesperada para risco de crédito e, por conseguinte, estimar o capital mínimo regulamentar.

Desse modo estabelecendo tentativa de proteção aos agentes credor referente ao risco incorrido em suas carteiras.

De acordo com o Relatório BIS (2006), as instituições financeiras devem estimar a probabilidade de *default* na IRB básica e utilizar os parâmetros sugeridos pela Basileia em relação à IRB Avançada, podendo utilizar seus modelos internos de risco, conforme sugere a regulamentação. Porém o Acordo ainda sugere que a ideia principal é que cada instituição possa construir seu próprio modelo de risco.

Thomas *et al.* (2005) mencionam que a falta de modelos de classificação de risco de crédito para carteiras de empréstimos é realmente um fator relevante para o Novo Acordo da Basileia e explicam que há somente indicações dos reguladores referentes à forma dos modelos, como por exemplo, os modelos do tipo Merton (1974), semelhante ao modelo *Credit Metrics*, JP Morgan (1997) no qual os valores de correlação são atrelados a modelos de classificação capazes de produzir diferentes cenários de risco.

Desse modo justifica-se a expectativa do Acordo para que as instituições financeiras construam seus modelos baseados na gestão do seu risco interno.

Gurtler e Heithecker (2005) mencionam que tais modelos que analisam o valor da carteira em um determinado horizonte de tempo, por exemplo, 12 meses, sendo eles modelos do tipo Merton (Merton, 1974) ou o modelo unifatorial de Vasicek (Vasicek, 2002) são os modelos escolhidos para fundamentar a adequação de capital regulamentar (Basileia II) e adotado pelo Comitê BIS (2004).

Porém Wilson (1998) argumenta que mesmo com modelos adotados pela regulamentação, há dificuldades na modelagem de risco de crédito em carteiras, principalmente pelo desconhecimento em relação ao risco individual nas carteiras de crédito e do risco de crédito atrelado às maturidades de um empréstimo, *Maturity (MT)*, (GURTLER E HEITHECKER, 2005).

Porém ainda Bellotti e Crook (2008), inspirados pelo Novo Acordo de Capital da Basileia II, propuseram o modelo *Loss Given Default (LGD)* com dados ingleses de cartões de crédito para o mercado de varejo e investigaram a perda incorrida pelas instituições financeiras a partir da inadimplência, possibilitando a estimativa da Perda Esperada (PD) para empréstimos individuais.

Entretanto Malik e Thomas (2009) mencionam que a medição da perda esperada, PD, bem como a previsão de *default* se realiza em função de se estimar o capital mínimo referente à carteira de empréstimos ao consumidor e que não será útil somente no nível individual, mas também entre os diversos segmentos da carteira, além das possibilidades de testes de estresse em carteiras de empréstimos pelas instituições.

Desse modo, Thomas *et al.* (2005) argumentam que a abertura feita pelo Novo Acordo BIS (2007) para que as instituições financeiras desenvolvam seus próprios modelos de risco de crédito é necessário que estes absorvam o impacto do Acordo como um todo e que possam também explicitar os fatores de risco para operações de longo prazo, haja vista a deficiência de dados históricos creditícios das instituições.

Por fim Malik *et al.* (2009) informam que o Acordo da Basileia (BCBS, 2004), nos últimos anos foi o elemento motivador para o desenvolvimento de modelos de risco de crédito, ainda que não se estabelecem critérios claros e orientações sobre tema .

Jackson (1999) considera que o impacto da Basileia na modelagem preditiva de risco de crédito centra-se principalmente na adoção de requisitos mínimos exigidos para alocação de capital, sendo este o ponto crucial do Acordo.

2.2.1 Regras do Banco Central do Brasil relacionadas a crédito

No Brasil, a Resolução CMN 2.099/94 e a Resolução 2682/99 instituíram o Acordo da Basileia no mercado de crédito brasileiro e informam a obrigatoriedade dos bancos em atender aos requerimentos estabelecidos pelos reguladores internacionais, bem como os valores mínimos de capital realizado, os valores do patrimônio líquido ajustado pelo grau de risco das instituições, bem como a classificação e provisão de empréstimos.

A Resolução CMN 2.099/94 estabelecia que o índice de capital de risco ponderado exigido a todos os bancos brasileiros fosse de 11% dos ativos com base no risco ponderado conforme Circular n º2.784, (Banco Central do Brasil, 1997). Desta forma, em virtude da adequação do capital, o Banco Central do Brasil exigia, em linhas gerais, que as instituições financeiras fossem submetidas à regulamentação nacional, de forma similar as do Acordo da Basileia em relação à adequação de capital com exceção, por exemplo, do patamar mínimo de capital de 11%, em vez de 8% conforme exigido pelo Acordo.

Entretanto, o Banco Central do Brasil anunciou em 2011 o aumento do índice de 11,0% para 16,5%, para o capital mínimo requerido para as instituições financeiras em relação às operações de crédito a pessoa física, com prazo de empréstimos superior a 24 meses, como

medida preventiva em relação ao crescimento do mercado de crédito ao consumo e do endividamento da base de clientes tomadores de empréstimo para o consumo, (Banco Central do Brasil, 2011).

A circular do Banco Central 3.498/10 aumentou a exigência de capital para operações de crédito de longo prazo, destinadas ao consumo, considerando a questão de risco de crédito nas operações de crédito a pessoas físicas, justificando que o aumento das exigências em relação às operações de crédito é a forte expansão do crédito sustentado pelo aumento da base de clientes e pelo alongamento dos prazos dos empréstimos.

Desse modo, recentemente, outra medida em relação à avaliação do risco para o segmento de empréstimo ao consumo foi adotada pelo governo brasileiro, o Cadastro Positivo por meio da MP 518/10, com o objetivo de mitigar o risco originado pelo aumento do índice de endividamento das famílias.

O Banco Central do Brasil, por meio do Comunicado 12.746/04, demonstra que o objetivo em relação à adoção do Novo Acordo é adaptar as regras requeridas à realidade do setor bancário brasileiro, de forma gradual.

Posteriormente, os comunicados 16.137/07 e 19.028/09 divulgaram as regras para implementação de Basiléia II no Brasil, bem como os critérios a serem adequados em relação ao nível de riscos associados às operações em relação ao requerimento de capital regulamentar.

Como procedimento, o Banco Central do Brasil estabeleceu um cronograma para implementação no mercado brasileiro, aplicado a todas as instituições do sistema financeiro nacional.

A Tabela 1 sintetiza a evolução das adequações das regras da Basiléia no contexto brasileiro.

Tabela 1: Cronograma de adequação as regras da Basileia II

2005	Revisão dos requerimentos para adoção da abordagem simplificada;
2007	Estabelecimento dos critérios de elegibilidade para a implementação da abordagem baseada em classificações internas para risco de crédito;
2008 2009	Estabelecimento de cronograma de validação da abordagem baseada em classificações internas para risco de crédito;
2009 2010	Validação dos sistemas de classificação interna pela abordagem avançada para risco de crédito;
2010	Estabelecer critérios de elegibilidade para a implementação da abordagem baseada em classificações internas, e divulgação do processo de solicitação de autorização para uso da abordagem baseada em classificações internas para apuração de requerimento de capital para risco de crédito e;
2012	Início do processo de autorização para uso das abordagens básica e avançada baseadas em classificações internas para apuração de requerimento de capital para risco de crédito.

O comunicado 19.028/09 relaciona ainda, a aplicação do Acordo no Brasil referente, essencialmente Pilar I, e ações correspondentes aos pilares II e III do Acordo da Basileia II e propõe o cronograma a seguir para alinhamento e implementações, visando o fortalecimento das medidas já adotadas, propondo maior solidez do mercado financeiro.

2.3 Formas de análise em modelos de classificação de risco de crédito

Historicamente, o método de *Credit Scoring* desenvolvido por Fisher (1936) e que, segundo Thomas *et al.* (2003), foi o trabalho pioneiro em classificação de risco de crédito concebido por duas categorias: Modelos de Concessão (*application scoring*) e Modelos Comportamentais (*behavioural scoring*), ambos considerando a estatística como técnica-base padrão.

Para Thomas (2003), os próximos 50 anos marcarão o desenvolvimento de modelos e sistemas de classificação de risco de crédito que possam integrar as decisões de crédito em nível estratégico de uma empresa.

Porém Mester (1997) menciona que a necessidade do desenvolvimento de sistemas de classificação de risco de crédito não é de agora, adverte, entretanto que tal necessidade já era sentida nos últimos 10 anos e, portanto alerta que tais modelos ou sistemas devem ser capazes de eliminar ou reduzir efeitos de dados históricos em relação aos novos clientes, prover sistemas capazes de estimar a probabilidade de default com precisão sem perder a capacidade preditiva em relação a amostras de grande volume.

Mais recentemente Malik e Thomas (2009), mencionam que em relação aos modelos que possam fornecer capacidade preditiva a abordagem de scorecards também utilizada nos últimos anos não considera o comportamento de crédito futuro do cliente, apenas um momento específico da vida creditícia do cliente e, assim não concebe as mudanças nas condições econômicas de mercado e da qualidade dos empréstimos ao longo do tempo.

Desse modo Mester (1997) ainda argumenta que as verdadeiras limitações na utilização de *credit score* pelas empresas ou instituições financeiras são as reduções de custos em detrimento ao processo de concessão de crédito, pois poderão retornar como despesas, caso haja deficiência ou baixo desempenho na classificação de risco durante o período analisado ou da origem do empréstimo.

Hand e Henley (1997) afirmam que há diversos modelos e técnicas estatísticas, que associadas, podem produzir modelos únicos, por exemplo, métodos matemáticos ou mais sofisticados como o modelo Markov multiestado, (JACKSON, 2006), ou Redes Neurais Artificiais de (RIPLEY, 1994), ou ainda o modelo de classificação direta e indireta de (Li e Hand, 2002). Mas na visão de Caouette et al. (1998, p.40) só há um único caminho para as instituições financeiras se protegerem contra as perdas em relação à concessão de crédito: a adoção de novas abordagens e técnicas mais sofisticadas em relação às utilizadas até o momento.

2.3.1 Modelos de concessão de crédito

Fantazzini e Figini (2009) mencionam que Altman foi o primeiro pesquisador a utilizar modelos estatísticos para estimar a probabilidade de inadimplência das empresas, com o uso

da variável que se tornou conhecida como Z-Score (1968) e Zeta (1977), em modelo padrão discriminante.

Foi por muitos anos um dos modelos mais importantes para o cálculo de risco e depois foram substituídas pelos modelos de regressão logística, redes neurais, sistemas computacionais entre outros. Estes modelos são amplamente utilizados para aplicação empírica e teórica na área de risco de crédito.

Altman (1968) desenvolveu a teoria em que os índices, se analisados em um cenário multivariado, poderiam ter maior significância do que a técnica comum de comparação de índices financeiros utilizados. Decorre daí a abordagem da Análise Discriminante como sendo uma técnica estatística utilizada para classificar um dado dentro de uma série de grupos dependentes de características individuais das observações.

Assim Altman *et al.* (1977) confirmam suas observações com o modelo de análise discriminante, prevendo corretamente a falência em 94% das amostras iniciais com 95% de todas as empresas classificadas como falidas no período de 1969 a 1975, utilizando medidas de falência relacionadas ao mercado na previsão.

E posteriormente Altman (2000) amplia as discussões sobre previsão de inadimplência, argumentando a importância da identificação das características específicas de cada negócio, principalmente suas falhas, a fim de especificar e quantificar as variáveis que são indicadores eficazes e possíveis previsores de risco corporativo.

Porém foi somente em Altman (2002) ao analisar os modelos de *Scoring* de crédito, bem como os modelos KMV / EDF de Merton (1974) em virtude da Basiléia II, referente ao capital em ativos de crédito e às taxas de inadimplências, e, portanto o autor concluiu que as ferramentas propostas pelos modelos para evitar as perdas ocorridas não são suficientes, é preciso que, além de modelos previsores, haja também cultura de risco nas empresas.

Malik e Thomas (2009) concordam com Altman (2002) e consideram que o desenvolvimento de modelos robustos, além dos já utilizados para as carteiras de empréstimos ao consumidor, é importante e se configura como consequência do fracasso dos modelos de classificação de

risco das agências de *rating* de crédito em definir corretamente o valor dos títulos lastreados pelas hipotecas durante a crise de crédito que ocorreu em 2007 e 2008.

Por fim Thomas e Hand (2005) mencionam que os modelos previsores de risco de crédito mais utilizados nos últimos anos em relação à classificação de risco são os modelos de concessão de crédito com tratamento estatístico de Regressão Logística e justifica que tais modelos ainda são os mais utilizados por demonstrarem mais robustez em relação a outras técnicas baseadas em análise multivariada, ou análise discriminante, como visto em Altman (1968), e por último afirmam que o desenvolvimento de modelos de risco de crédito será de grande importância em termos de proteção ao sistema financeiro como vantagem competitiva entre os bancos.

Em relação a novos modelos, Ripley (1994) menciona que desde 1985, a técnica estatística de Redes Neurais tem sido desenvolvida na tentativa de capturar fatores flexíveis que possam indicar o comportamento futuro de um tomador de crédito e que apesar da técnica ter sido desenvolvida há mais de 20 anos, esta vem sofrendo diversas revisões.

Por exemplo, atualmente, o modelo bayesiano (BAESENS *et al.*, 2005) sob uma nova ordem de modelos neurais é exemplo de uma nova tendência nos modelos de riscos; os quais, conforme os autores consideram que a dificuldade encontrada na modelagem de dados de crédito ao consumo se realize em função da dificuldade em capturar a previsão do padrão da carteira de empréstimo e do pagamento antecipado, observados em uma amostra, o que vem dificultando o desenvolvimento de novos modelos.

Dificuldades não faltam no campo de desenvolvimento de modelos de risco de crédito, Verstraeten e Poel (2005) mencionam o viés de amostragem como uma dificuldade importante na construção dos modelos, principalmente em função do grande volume de dados, que é uma limitação bastante séria do ponto de vista estatístico. Entretanto Fantazzini e Figini (2009) argumentam que modelos mais flexíveis ou combinados podem diminuir tais limitações, como, por exemplo, modelos híbridos de risco de crédito.

Assim nos últimos 10 anos percebe-se também a utilização do modelo de Mineração de Dados (HIAN e CHWEE, 2006) como um algoritmo de aprendizagem com a expectativa de

extrair e evidenciar padrões repetitivos em uma determinada amostra e, assim, mapear o comportamento de um tomador.

2.3.2 Modelos de concessão: principais modelos de carteira de crédito

Em relação aos modelos de carteira de crédito, as instituições financeiras devem atender aos requisitos mínimos exigidos pelos reguladores em relação a empréstimos ao consumidor e deverão considerar principalmente a segmentação adotada, produto e classificação dos mutuários.

Wilson (1998) alerta que antes de qualquer esforço aplicado ao desenvolvimento de modelos de carteira previsores de *default*, é preciso que se defina primeiramente o conceito de risco e como diferentes cenários macroeconômicos podem afetar o comportamento do indivíduo em relação ao risco total de uma carteira de crédito, e assim, determinar um padrão.

Thomas *et al.* (2005) afirmam que as instituições deverão analisar cinco anos de dados de crédito, além de calcularem os parâmetros essenciais para qualquer segmento de crédito adotado e conseqüentemente haverá dificuldade no desenvolvimento de modelos de classificação de risco para carteiras de crédito devido ao grande volume de dados e à falta de dados históricos creditícios pelas instituições.

Tal dificuldade é vista por Wilson (1998) em forma de advertência, pois é necessário que se defina primeiramente o risco em carteiras de crédito e como diferente cenário macroeconômico (considerando indústria e segmentações) possam afetar o comportamento de um indivíduo em relação a um portfólio.

Para Vasicek (2002) o risco individual em carteiras de varejo pode ser avaliado em função da distribuição das perdas futuras dos empréstimos de uma instituição conforme modelo de Vasicek (VASICEK, 2002) e posterior agrupamento por produto, segmento e classificação dos dados da amostra em setor econômico, país e *rating*.

Já Crouhy *et al.* (2001) mencionam que os modelos de risco de portfólio mais difundidos no mercado são: *Creditmetrics*, do banco J. P. Morgan (1997); *CreditRisk+*, da Credit Suisse

Financial Products (CFSF, 1997); *Credit PortfolioView*, da McKinsey, (Wilson, 1997) e o KMV, (KMV, 1993).

Saunders (1999) explica que o surgimento de novos modelos de classificação de risco para carteiras de crédito é resultado da aplicação da teoria de opções para avaliação de risco de empréstimos e *bonds*, por exemplo, o modelo (*Credit monitor Model*), que produz e atualiza previsões de *defaults* para quase todas as empresas e bancos que possuem ações negociadas no mercado, assim como o modelo estrutural KMV, derivado da teoria moderna de finanças, na qual o capital de uma empresa é considerado uma opção de *call*.

Dessa forma Thomas (2009) concorda e acrescenta que o modelo KMV estrutural de risco de crédito corporativo desenvolvido por Merton (MERTON, 1974) assume que uma empresa entra em inadimplência quando o valor de seus débitos ultrapassarem a mesma proporção de seus ativos.

Segundo Gürtler e Heithecker (2005), o desenvolvimento de modelos para classificação de risco em carteiras de crédito tem evoluído significativamente, sendo os mais populares do tipo Merton (1974) e os unifatoriais de Vasicek (2002), que constroem o fundamento utilizado para adequação de capital conforme a Basileia II, adotada pelo Comitê de Supervisão Bancária (2004).

Porém Vasicek (2002) argumenta que o capital necessário para suportar uma carteira de títulos de dívida depende da distribuição de probabilidade da perda, uma vez que a qualidade de crédito do credor vai depender da probabilidade da perda da carteira em exceder o capital próprio, confirmando os argumentos de Wilson (1998), que sugere conhecer o risco individual de cada tomador de crédito para se definir o risco de portfólio.

Mas a preocupação para Galindo e Tamayo (2000) ainda permanece em encontrar a correta estimativa ou a variável preditiva de risco individual em portfólios de crédito de instituições financeiras.

Entretanto Gürtler e Heithecker (2005) mencionam que tais modelos consideram apenas as perdas devidas à inadimplência de empréstimos e, portanto, com limitações quanto à maturidade dos empréstimos, *Maturity* (MT). E em relação à maturidade dos empréstimos,

Vasicek (2002) complementa que a distribuição de probabilidade em limitar as perdas de carteira foi derivada sob a suposição de que todos os empréstimos de carteira possuem a mesma maturidade, a mesma probabilidade de *default*, e a mesma correlação *pairwise* dos mutuários.

Porém impulsionados com o surgimento de produtos inovadores de dívida corporativa Bharath e Shumway (2008) propõem examinar a precisão e a contribuição do modelo padrão, Distância de Merton (DD) comparando os preços de *Bonds* na forma funcional em uma alternativa *naive*, denominada “ingênua” para o modelo de Merton (1974), entretanto, o modelo não produziu resultado estatístico suficiente para estimar a probabilidade de *default*, apesar de sua forma funcional ser útil para a previsão de padrões.

Apesar da forma funcional dos modelos de portfólio para Banasik et al. (1999) mencionam que atribuir nota de classificação de risco, estimada pela previsão do padrão recorrente em relação ao tempo dos dados analisados, ou seja, modelos que possam estimar a possibilidade futura de mudança no padrão são a grande expectativa das instituições em relação à provisão de recursos financeiros e subsídios para a tomada de decisão em relação a risco de crédito.

De acordo com Galindo e Tamayo (2000) a possibilidade de estimar mudança no padrão de consumo em uma determinada carteira de crédito é importante e caracteriza que o ciclo de negócio é afetado por correlações dos fatores de risco macroeconômico e, portanto a interpretação destes fatores poderá levar a previsões mais precisas de perda e promover redução considerável em relação ao capital econômico, conforme adverte Jacobs (2011).

Deste modo Rösch e Scheule (2004) sugerem, portanto, que o modelo de portfólio para concessão de crédito pode ser concebido com variáveis macroeconômicas defasadas de fatores de risco, explicando o risco de crédito incorrido nas carteiras de crédito no varejo.

E os modelos *Creditmetrics* (1997) e *Credit Portfólio View* (1997), segundo Gürtler e Heithecker (2005), são modelos de gestão de portfólio que tem como objetivo proteção às instituições financeiras à exposição de risco excessivo no mercado.

Wilson (1998), por sua vez, acrescenta que especificamente no modelo *Credit Portfolio View* a probabilidade de *default* está condicionada às variáveis macroeconômicas. Assim como

Saunders (1999) que menciona que o *Credit Portfolio View* reúne um conjunto de fatores macroeconômicos e cenários distribuídos que possibilitam a leitura subjacente da economia, onde a probabilidade de *default* é uma função logística e as variáveis macroeconômicas, uma probabilidade para as transições das matrizes de risco.

Saunders (1999) ainda afirma que, diferentemente do *Credit Risk Plus*, a probabilidade de cada empréstimo ir a *default* é vista como uma variável de acordo com a Distribuição de Poisson.

Para Crouhy *et al.* (2000), o *CreditRisk+* (1997) da *Credit Suisse Financial Products*, ao contrário do KMV (1993), estabelece que o risco de inadimplência não esteja relacionado à estrutura de capital da empresa, diferenciando-se em sua estrutura analítica.

Porém Banasik *et al.* (1999) propõem a pergunta de “quando” os mutuários irão se tornar inadimplentes?, E não somente se foram inadimplentes. A resposta para esta pergunta, segundo os autores é parte fundamental no desenvolvimento de modelos de risco de crédito para o consumo, sendo possível a partir desta resposta, estimar um padrão de “tempo de *default*” em uma determinada carteira.

Para Malik e Thomas (2009) o modelo de risco de crédito para portfólios está diretamente relacionado à exigência do Acordo da Basiléia II (BCBS, 2004), e alerta que não é suficiente apenas classificar corretamente os clientes de acordo com seu risco de inadimplência, é necessário também medir com precisão a perda esperada (PD), pois são informações utilizadas para calcular o capital mínimo necessário referente à carteira de empréstimos ao consumidor.

Porém para Bardos (2007) não é somente a previsão da perda esperada (PD) a necessidade nos modelos de portfólios o autor sublinha a importância em relação à perda inesperada, ou seja, os quantis extremos, em especial, para os clientes de bancos comerciais, contexto em que é necessário estimar a probabilidade de *default* de cada empresa em um determinado horizonte de tempo.

Vasicek (2002) ainda alerta que, em relação à perda, é necessário caracterizá-la corretamente, pois pode haver perdas realizadas em empréstimos vencidos ou uma desvalorização do valor

de mercado dos empréstimos em detrimento do prazo e do período dos empréstimos em relação a uma determinada data futura.

Os modelos de portfólio analisados por Bellotti e Crook (2008), modelos de *Loss Given Default* (LGD) testaram variáveis relacionadas ao nível de atividade econômica com o objetivo de investigar a perda incorrida pelas instituições financeiras a partir da inadimplência, uma vez que o *Loss Given Default* possibilitaria às instituições avaliarem seus riscos em suas carteiras de crédito, e também, para fins da implementação do Novo Acordo de Capital (Basileia II).

Ainda em relação aos modelos de carteiras de crédito, Jacobs e Karagozoglu (2010) informam que a previsão de LGD de títulos e empréstimos é importante para o atual estágio em relação a risco de crédito e mencionam que a informação a partir de qualquer patrimônio ou a mercados de dívida em torno do tempo de padrão (medidas de endividamento ou preços ou retornos de capital acumulado, respectivamente) têm poder preditivo em relação ao LGD final.

Jacobs (2010) utilizou uma amostra de empresas inadimplentes classificadas pela *Moody's* em um modelo preditivo econométrico de regressão múltipla no intuito de encontrar determinantes que possam mensurar o risco de EAD, Exposição ao *Default*, e descobriram que o risco pode ser reduzido através de alavancagem e maior liquidez. Em contrapartida o risco poderá aumentar devido ao tamanho da empresa, classificação da empresa, e pelas dívidas bancárias de acordo com a estrutura de capital do devedor, por fim identificaram uma relação inversa entre risco de EAD e inadimplência, ou seja, o risco de EAD pode diminuir em relação à inadimplência.

Para Thomas (2003), a questão principal é em relação à validação e verificação das estimativas de probabilidade de *default* e perda esperada, dado o não cumprimento (LGD) e ainda que tais estimativas devam ser a média de longo prazo da PD e LGD para o período de 12 meses. Neste sentido, na modelagem seria preciso incluir os ajustes em relação às mudanças na economia e condições de mercado, consideração também acordada por Jacobs (2010) em sua abordagem sobre capital econômico.

E por fim Gürtler e Heithecker (2005) mencionam que os modelos *Creditmetrics* e *Credit Portfolio View* são fundamentados no conceito de marcação a mercado (*fair value*), que o diferencia, segundo Saunders (1999), dos modelos *Creditmetrics* e KMV, que são analíticos, como no modelo tipo Merton, no qual o valor do capital de uma empresa e a volatilidade são indicadores de risco e previsores de *default* (*drivers*), como em Rösch e Scheule (2004).

2.4 Estudos empíricos sobre análise de crédito para empréstimos ao consumo

Para os bancos, a primeira necessidade em relação a modelos de risco de crédito é poder generalizar, classificar e quantificar o risco inerente de cada tomador e sua capacidade preditiva de risco de *default*, bem como gerar *ratings* consistentes a uma determinada carteira de tomadores de crédito, conforme afirmam Caouete *et al.* (1998).

Por meio do modelo Cox de risco proporcional, Cox (1972) e Banasik *et al.* (1999) modelaram dados de clientes que foram classificados como inadimplentes, mas que deixaram de ser no presente (ou por terem sido pagos os empréstimos em atraso ou por morte). Os resultados foram competitivos para o período de 12 meses em relação à abordagem de regressão logística, entretanto, sugerem modelos mais sofisticados para amostras e prazo maiores.

Para Andreeva *et al.* (2003), o intuito é desenvolver modelos de risco de crédito estratégicos adequados à área de Marketing, utilizando Análise de Sobrevivência em modelo de risco proporcional. Para isso, analisaram dados de uma companhia de seguros belga para estimar a probabilidade padrão de previsão do tempo requerido pelo cliente em realizar um seguro e concluíram que o modelo em questão fornece fundamentos para estimar a saúde financeira do tomador de crédito.

Rösch e Scheule (2004) consideram que os parâmetros de probabilidade de incumprimento podem ser modelados utilizando *drivers* defasados de risco macroeconômico para explicar o risco de crédito, no qual o segmento varejo esteja exposto.

E em relação a empresas não financeiras, Fantazzini e Figini (2009) observaram dados anuais de 1996 a 2004 para 1.003 pequenas e médias empresas (PME) da principal agência de

classificação da Alemanha, com abordagem não paramétrica, baseada em *Forests Survival Random (RSF)*. Os autores compararam o desempenho com um modelo *logit* padrão e concluíram que o modelo *logit* superou um modelo mais complexo RSF em testes fora da amostra, em relação à precisão das previsões, embora o modelo RSF tenha proporcionado uma melhor descrição das PME em relação aos dados da amostra, conforme autores sugerem que o problema de viés de amostra possa ser melhorado em pesquisas futuras.

Em relação à *Loss Given Default (LGD)*, Bellotti e Crook (2008) utilizaram um modelo com técnicas de modelo de regressão, árvore de decisão, em conjunto a variáveis relacionadas ao nível de atividade econômica para prever perdas em cartões de crédito para o mercado de varejo britânico a partir da inadimplência. Os resultados confirmaram que o modelo é capaz de possibilitar avaliações de risco em carteira de crédito, bem como perdas esperadas para empréstimos.

Por sua vez Jacobs e Karagozoglou (2010) modelaram uma amostra de inadimplentes dos EUA no período 1985 a 2006 e demonstraram que a informação a partir de qualquer patrimônio ou mercados de dívida em torno do tempo de padrão (medidas de endividamento ou preços ou retornos de capital acumulado, respectivamente) tem poder preditivo em relação ao LGD final.

Assim, considera-se que o interesse em pesquisas em análise de crédito para empréstimos ao consumo é recorrente e necessário.

2.4.1 Empréstimos por meio de cartão de crédito

Thomas et al. (2011) mencionam que o advento de cartões de crédito surgiu na década de 1960 e transformou o meio de compra dos consumidores, possibilitando financiamento de todas as suas compras, de grampos de cabelo até viagens de férias apenas utilizando cartões. Segundo o autor, a expansão do crédito via cartões de crédito é histórica e relata que, em 2002, havia mais de 500 milhões de cartões de crédito na Europa e que cerca de 70% dos pagamentos realizados pela internet é por meio de cartão de crédito.

No Brasil, conforme o Relatório do Banco Central do Brasil, o estoque de cartões de crédito emitidos passou de 39 milhões em 2003 para 118 milhões em 2007, crescendo aproximadamente 200% no período, ou seja, os cartões ativos aumentaram 190%, atingindo a marca de 66,6 milhões de cartões de crédito. Consequentemente, a quantidade de transações com cartão de crédito vem crescendo consistentemente desde 2002, passando de 195 milhões para 603 milhões em 2007, representando o crescimento de 209% no período (Banco Central do Brasil, 2010).

No mercado de crédito brasileiro, o Banco Central do Brasil (2011) anuncia que o saldo das operações de crédito no Brasil alcançou a marca de R\$ 1,5 trilhão ao final de 2010, representando 45,5% dos ativos do sistema bancário, correspondendo à expansão de crédito referente a 21,7% em relação a dezembro de 2009, influenciado, segundo o relatório, (Banco Central do Brasil, 2011), pela melhora na expectativa dos consumidores e das instituições em relação à oferta de crédito. Deste modo, conforme o Relatório de Cartões de Crédito (Banco Central do Brasil, 2010) apresenta que a participação do crédito ao consumo não se realiza somente por meio de empréstimos tradicionais.

Nos últimos anos percebe-se que, o segmento de cartões de crédito, é responsável por cerca de 60% das transações de crédito, com o estoque de 118 milhões de cartões de crédito emitidos entre 2003 e 2007, crescendo aproximadamente 200% nesse período, o que demonstra forte preocupação por parte dos agentes reguladores sobre este tipo de segmento de crédito.

Segundo a Associação Brasileira de Cartões de Crédito (2011), o mercado de cartões acompanhou o crescimento da economia brasileira, movimentando 1,8 milhões de transações, com um faturamento de 23% de aumento em relação ao mesmo período em 2010. Essa evolução é reflexa da incorporação do hábito do uso de cartão de crédito nas classes C e D e entre as pessoas com mais de 60 anos, acompanhando o crescimento do acesso à conta bancária, conforme o estudo da Associação Brasileira de Cartões de Crédito (ABCS, 2010).

No entanto, o Banco Central do Brasil (2010) vem incentivando o uso racional do cartão de crédito com o intuito de proteger a economia contra o superendividamento, que pode incorrer em fonte de riscos operacionais para as instituições integrantes do Sistema Financeiro Nacional (SFN). E elevou, para isso, o requerimento mínimo de capital para risco de crédito

para 150% como fator de ponderação de risco (FPR) aplicado ao cálculo do Patrimônio de Referência Exigido (PRE) das exposições relativas a operações de crédito e de arrendamentos mercantis financeiros contratados por pessoas físicas.

Porém a Fecomércio, Federação do Comércio de Bens e Serviços do Estado de São Paulo (2011), por meio do índice de endividamento e inadimplência do consumidor através da Pesquisa de Endividamento e Inadimplência do Consumidor (PEIC) indica que a maior parte das dívidas assumidas pelos consumidores brasileiros é realizada pelo segmento de cartões de crédito, responsável por 68% do endividamento dos brasileiros somente no estado de São Paulo. O indicador Serasa Experian registra prejuízo de 25% em 2010 no empréstimo rotativo de cartão de crédito comparado com 2009.

Porém, o Relatório do Banco Central do Brasil (2011), adverte que o valor financiado no crédito rotativo vem aumentando em relação ao valor transacionado, sendo o prazo médio de recebimento pela venda realizada com cartão de crédito pelos estabelecimentos de 30 dias, diferentemente do cenário internacional, em que este prazo é de dois dias em média (Banco Central do Brasil, 2010).

O Banco Central ainda informa que tal prática isenta os emissores do custo de oportunidade do período sem cobrança de juros, dado que em média, eles recebem em 28 dias, podendo gerar receita de *float*.

Porém, segundo a Associação Brasileira das Empresas de Cartões de Crédito e Serviços (Abecs) em 2011 a expansão do faturamento dos cartões de crédito foi resultado do aumento dos gastos dos brasileiros no exterior, segundo revela a pesquisa da Abecs (Associação Brasileira das Empresas de Cartões de Crédito e Serviços). As compras feitas com cartões de crédito em outros países movimentaram R\$ 21,2 bilhões em 2011, o que representa um crescimento de 19% em relação a 2010.

Portanto a expectativa do crescimento do mercado de cartões de crédito, segundo a Associação é que o mercado de cartões de crédito movimente cerca de R\$ 650 bilhões em pagamentos, uma expansão de 20% na comparação com 2010, conforme estimativas preliminares da Associação Brasileira das Empresas de Cartões de Crédito e Serviços (Abecs).

2.4.2 Variáveis explicativas utilizadas na modelagem de risco de crédito

Diversos autores (conforme apresentado na Tabela 2) demonstraram que a inclusão de variáveis macroeconômicas, por exemplo, taxa de juros ou índice de desemprego, entre outros, em modelos de risco de crédito, podem melhorar a eficiência dos modelos, possibilitando a estimativa da probabilidade de *default* em uma determinada amostra de crédito.

Tabela 2: Variáveis Macroeconômicas utilizadas na modelagem de risco de crédito para consumo

Autores / Ano	Variáveis do Indivíduo	Modelo	Base de Dados	Variáveis Macroeconômicas
Peasaran et al. (1946)	Círculo de negócio doméstico e externo; especificidade da empresa; critério para <i>default</i> e <i>ratings</i> de crédito.	Modelo de risco do tipo Merton com variáveis macroeconômicas		Preço de combustível (<i>oil gás</i>) e preço das ações (<i>equity prices</i>)
Thomas et al. (2001)	Classificação Comportamental (<i>behavioural score</i>)	Dois modelos: Modelo Markov Análise de Sobrevivência	Carteira de Empréstimos ao consumo	Taxa de juro atual
Rosch e Scheule (2004)	Juros anuais dos 5 principais Bancos Comerciais Americanos; Casa própria; cartão de crédito e outros tipos de empréstimos individuais.	Modelo <i>Credit Metrics model</i>	Demonstrações financeiras trimestrais, salário e renda bancária norte-americana.	Varição no índice de preço ao consumidor (Inflação); taxa interbancária; taxa de juros; variação do PIB e variação da produção industrial.
Thomas et al. (2005)	Identificação do cliente, distância entre compras realizadas e tempo de relacionamento, gênero, situação financeira, <i>scoring</i> de crédito de pagamentos efetuados.	Modelo Cox de risco proporcional em Análises de Sobrevivência	Empresa Internacional de Seguros individuais, maior seguradora na Inglaterra, média de idade dos clientes de 0 a 100 anos.	Riqueza, (<i>Proxy</i> utilizada pelo índice de valorização imobiliária), Índice FTSE – todas as ações; renda média da população, Índice de confiança do consumidor e taxa de juros.

Tang et al. (2006)	Gênero, idade, e situação financeira atual. Fatores de risco econômicos e sócio-demográficos	Modelo de risco proporcional Cox e Weibull	Empresa Inglesa Internacional de Seguros 50.000 clientes (24.797 homens e 24.977 mulheres), período de 1999 até 2003.	Cinco variáveis econômicas externas; variáveis exógenas; Preço-Consumo; Índice de Confiança do Consumidor, Índice de Desemprego, Índice (<i>all shares</i>) e Taxa de Juros.
Bellotti e Crook (2007)	Renda, idade, casa própria, emprego, situação creditícia no <i>bureau</i> no momento da concessão de crédito.	Análise de Sobrevivência e Regressão Logística	100 mil contas bancárias de cartão de crédito de 1997 a 2005, dados ingleses.	Taxa de Juros, Índice de Desemprego, Índice Imobiliário, Renda, índice FTSE e de Produtividade.
Bellotti e Crook (2009)	Tempo de conta-corrente, renda e idade.	Modelo Tobit, Árvore de Decisão e Ordinary Least Squares (OLS).	55 mil contas de cartão de crédito Período de 1999 a 2005 dados em atraso, dados de Instituição Inglesa.	Taxa de Juros de bancos de varejo inglês, Índice de Desemprego na Inglaterra, e Renda.
Malik e Thomas (2009)	Prazo do empréstimo, classificação comportamental (<i>Behavioural scores</i>) e empréstimos concedidos.	Modelo de Risco Cox para previsão de PD	Dados mensais de um banco inglês de 2001 a 2005	Inflação no período de 01 ano; Variação nos preços (inflação); Variação interbancária da moeda; Retorno anual do Índice FTSE 100 e Variação trimestral do PIB.

De acordo com os autores citados acima, em sua maioria, relacionaram variáveis como, por exemplo, renda, tempo do relacionamento entre cliente e instituição, idade e casa própria como possíveis variáveis que possam indicar nível de sensibilidade do indivíduo em não honrar suas dívidas, e por outro lado, outros autores conforme demonstrado na Tabela 2 relacionaram taxas de juros básica do país, índices de desemprego, índices da cotação da bolsa de valores, índices de confiança do consumidor como possíveis *proxies* para a inadimplência.

Por exemplo, os resultados apontados por Pesaram *et al.* (2006) calcularam que os efeitos de um hipotético choque negativo nos preços de ativos asiáticos em relação à distribuição de perda em uma carteira de crédito demonstraram que devido ao choque, as perdas são

assimétricas e não proporcionais, refletindo a natureza altamente não linear do modelo de risco de crédito.

Bellotti e Crook (2007) e Jacobs e Karagozoglu (2010) utilizam em seus modelos variáveis macroeconômicas tais como taxa de desemprego, taxa de juros, PIB (Produto Interno Bruto) entre outras, com o objetivo de correlacionar o risco de *default* ou inadimplência relacionada à probabilidade de um evento de crédito ocorrer durante um determinado período de tempo em relação a uma carteira de empréstimos.

Malik e Thomas (2009) mencionam que a taxa de desemprego foi uma das variáveis significantes do modelo analisado e que, apesar de estar associada ao aumento do risco de inadimplência dos consumidores afetou de maneira geral as outras variáveis na equação de risco, embora atestem não haver nenhuma variação significativa na taxa de desemprego no Reino Unido durante o período analisado.

Poel e Lariviere (2003) modelaram dados de uma grande instituição financeira europeia em relação à incidência dos preditores de *Churn*, taxa de abandono do cliente, para demonstrar a probabilidade de clientes abandonarem o relacionamento com uma organização. Para aplicação em dados financeiros, entretanto, os autores mencionam que há limitações na construção de modelos como, por exemplo, dados sobre o comportamento do cliente, demográficos e do macroambiente.

Ainda Tang et al. (2007) no contexto da economia do Reino Unido, analisaram quatro variáveis macroeconômicas, que pudessem afetar o padrão de comportamento dos consumidores e que pudessem representar o clima geral econômico e de investimento. Segundo os autores, as variáveis escolhidas para o estudo são:

- (1) a variação percentual no índice de preços ao consumidor sobre 12 meses, sendo assim a intenção seria descobrir se a inflação sentida pelos clientes pudesse ter reflexo nos dados analisados e assim determinando que altos níveis do índice possam causar aumento na taxa de inadimplência de seus clientes;
- (2) média mensal da taxa de empréstimo interbancário, para que assim pudesse entender se o maior volume da moeda poderá corresponder à tensão geral na economia e à maior dificuldade em captar recursos para o pagamento da dívida;

- (3) o retorno anual sobre Log do FTSE 100: que poderia fornecer a ideia se a movimentação do mercado de ações poderá refletir e incentivar a compra de produtos financeiros e
- (4) o crescimento percentual do PIB (trimestral) em comparação com o mesmo trimestre no ano anterior da amostra.

Os autores justificam o uso de variáveis macroeconômicas em modelos de risco de crédito pela possibilidade de mensurar a influência macroeconômica no risco de *default* pelo tomador de crédito podendo até indicar diversas estratégias de negócios, inclusive de Marketing.

2.5 Inadimplência

A definição dada pelo artigo 580 do Código do Processo Civil Brasileiro considera como Inadimplente o devedor que não satisfaça a obrigação certa, líquida e exigível, consubstanciada em título executivo. (Redação dada pela Lei nº 11.382, de 2006).

O BCBS (*Basel Committee on Banking Supervision* – Comitê da Basileia sobre Supervisão Bancária, 2006, p. 100) define que inadimplência *default*, item 452, como sendo: o não pagamento em relação a um devedor específico quando o banco considerar improvável que o devedor pague na totalidade suas obrigações, sem que se tenha que recorrer à utilização de garantias (se possuir) e que o devedor esteja atrasado em mais de 90 dias em alguma obrigação material com o conglomerado financeiro.

Neste estudo, para fins de comparação de diferentes conceitos de inadimplência, é utilizado o conceito comum de inadimplente como “aquele que falta ao cumprimento de suas obrigações jurídicas no prazo estipulado,” neste estudo, 60 dias de atraso no pagamento da fatura de cartão de crédito.

Em relação ao Crédito Rotativo, modalidade de pagamento parcial da dívida, o pagamento do valor mínimo da fatura do cartão de crédito, hoje considerada 15% do valor total, conforme Banco Central do Brasil (2011), o que não se configura como inadimplência conforme contrato e entendimento entre as Administradoras de Cartões de Crédito no Brasil, pois é efetuado o pagamento mínimo da fatura do cartão de crédito.

Portanto neste estudo, onde se procura estimar a probabilidade de *default* de cada cliente ao longo de um determinado período, será utilizada a Regressão de Cox e suas variações. Considera-se o tempo de sobrevivência como o período de tempo observado no estudo até que os clientes atinjam o *status* de **0** para Censurado e **1**, Enquadrado, ou seja, inadimplentes. Os indivíduos que não atingiram esse nível de atraso, (60 dias) no período observado, 70 meses, foram considerados Censurados.

3 METODOLOGIA

3.1 Problemas de pesquisa

Através da técnica estatística de Análise de Sobrevida com variáveis macroeconômicas e variáveis de cadastro, aplicados aos modelos paramétricos e semiparamétricos é possível observar o aumento ou diminuição do risco de inadimplência em usuários de cartões de crédito?

3.2 Técnicas de análise

3.2.1 Análise de Sobrevida

A operacionalização desta pesquisa é parte primordial para o entendimento do conceito a ser entendido, que é o risco de *default* em contas de cartão de crédito. Assim, a escolha da pesquisa do tipo quantitativo, conforme Bryman (2004) deu-se pela dimensão escolhida do objeto a ser estudado, que é a relação entre risco e crédito de uma determinada amostra secundária de dados numéricos coletados com uma instituição financeira.

Creswell (2007) menciona que nas pesquisas quantitativas, as hipóteses e as questões de pesquisas são frequentemente baseadas em teorias que se deseja testar, portanto, o modelo escolhido concebe a ideia de explicar um fenômeno, *default* em conjunto com variáveis que indicam probabilidade de sensibilidade à hipótese deste estudo.

A Análise de Sobrevida é uma técnica estatística que analisa uma variável aleatória positiva. Tipicamente, essa variável é o tempo até a ocorrência de um determinado evento de interesse, também chamado de *tempo de sobrevivência*, Kleinbaum e Klein (2006).

Os precursores do uso da técnica de análise de sobrevivência na área médica foram Cox (1972) e Gil (1984) e seus estudos têm sido regularmente utilizados na área de finanças

modernas, onde se tem sido desenvolvido diversos modelos de classificação de risco de crédito, utilizando como tratamento estatístico de Análise de Sobrevivência, nas formas pontuais ou temporais. Para Bastos e Rocha (2006), a modelagem de tempo para dados de eventos é muito utilizada na área médica, bem como na engenharia estatística, onde neste contexto, a morte ou o fracasso é considerado um "evento", o que tradicionalmente na área médica apenas um único evento ocorre.

O tempo é uma variável importante para pesquisadores na área de Finanças e também para a área de risco. Thomas et al. (2005) mencionam que Narain (1992) foi o precursor nos estudos empíricos em análises de sobrevivência e técnicas estatísticas para classificação temporal de risco de crédito.

Stepanova e Thomas (2002) relatam que modelos que utilizam Análise de Sobrevivência têm sido desenvolvidos para classificação de risco de crédito por esta ser uma técnica flexível que permite inclusão de variáveis macroeconômicas e têm sido muito difundida nos últimos dez anos. Segundo o autor a construção de um modelo de *scorecard* apoiada em uma classificação temporal é melhorada quando aplicadas a base de análise de sobrevivência, em que o tempo é a previsão do evento de interesse.

Andreeva (2006) confirma que apesar dos diversos modelos e estudos na área de modelagem de risco de crédito, as aplicações de análises de sobrevivência iniciadas com os estudos de Narain (1992) têm demonstrado que estimativas de duração de tempo em relação ao prazo de empréstimos obtidos de um modelo exponencial podem significativamente melhorar as decisões de concessão de crédito aos tomadores.

Como por exemplo, Banasik *et al.* (1999), e Thomas e Stepanova (2002) utilizaram Análise de Sobrevivência em finanças por considerarem-na uma ferramenta altamente flexível para a construção de modelos de risco de crédito, principalmente pela possibilidade de inclusão de variáveis macroeconômicas.

Para Bellotti e Crook (2007) descrevem que as vantagens da Análise de Sobrevivência são: possibilidade de análise no processo de inadimplência do empréstimo, (ii) possibilidade de avaliar a probabilidade de rentabilidade de um empréstimo e (iii) fornecer estimativas de previsão de sobrevivência em função do tempo.

Malik e Thomas (2009) também utilizaram análises de sobrevivência para analisar possíveis padrões em carteiras de empréstimos pessoais em relação à probabilidade de *default*. Stepanova e Thomas (2002) em relação a análises de sobrevivência pesquisaram risco de crédito na modelagem de preços dos títulos (*Bonds*).

E ainda Andreeva (2006) analisou dados de crédito rotativo em cartões de crédito para realizar estimativas de duração de tempo de vida de um empréstimo obtidas por um modelo exponencial.

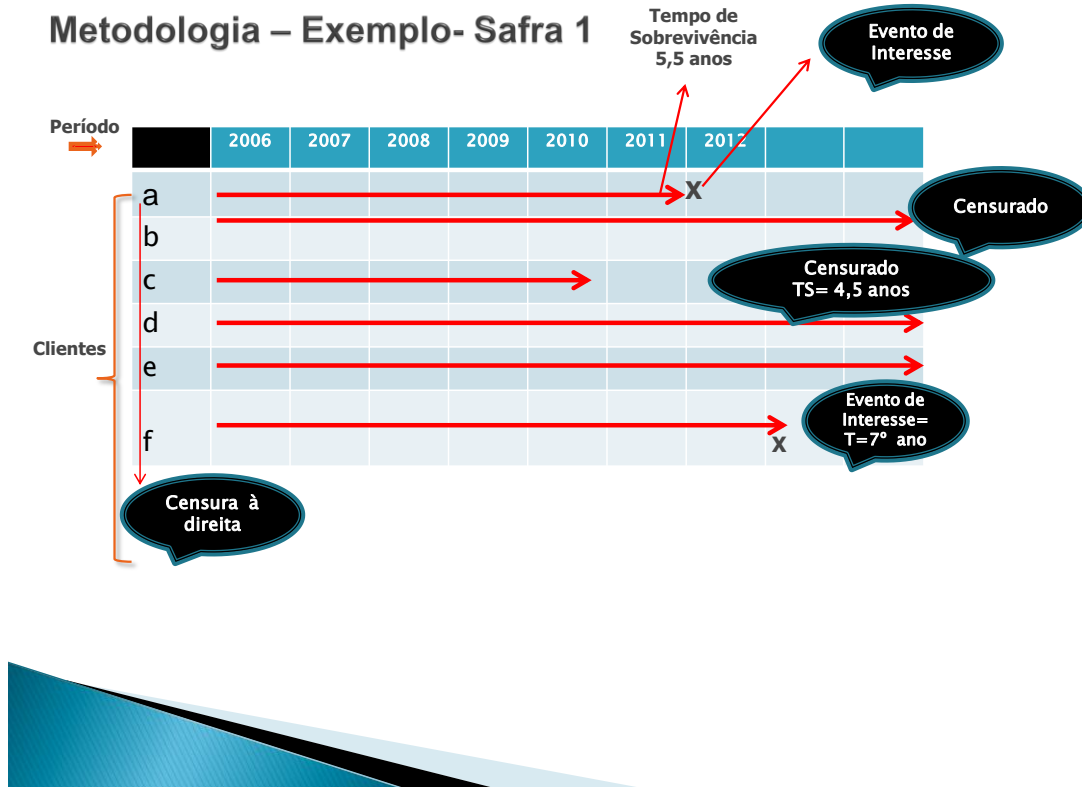
Porém Cox (1975) introduz um novo princípio para inferência, com base no conceito de risco parcial, que considera que cada indivíduo são valores disponíveis de uma ou mais variáveis explicativas. A função de risco é considerada como sendo uma função das variáveis explicativas e os coeficientes de regressão desconhecida, multiplicada por uma função arbitrária e desconhecida de tempo.

Tsiatis (1981) menciona que os modelos de regressão em análise de sobrevida, com observações censuradas, foram usados extensivamente nos últimos anos, principalmente o Modelo de Cox (1972) porque oferece forte consistência e normalidade assintótica para a estimativa de probabilidade máxima e parcial do parâmetro de regressão.

Gil (1984), entretanto, explica que não há exclusão nas abordagens, ou seja, o modelo não paramétrico pode indicar evidências não indicadas por um modelo paramétrico e vice-versa, sugere-se, portanto, o uso das duas abordagens para garantir resultados efetivos.

Os autores ainda mencionam que o contexto de estimação envolve a modelagem de tempo para dados de eventos modificados no decorrer do tempo, na qual a variável tempo é estudada até a ocorrência do evento de interesse, tempo de falha (*default/Inadimplência*), permitindo a observação de dados censurados, ou seja, quando o evento de interesse não ocorreu ou não existe possibilidade de verificação, conforme a Tabela 3:

Tabela 3: Tabela de Sobrevivência: representação de tempos de sobrevivência das transações de crédito



- O cliente *a* é seguido desde o início do estudo até ter ocorrido o evento de interesse no 5º ano; o tempo de sobrevivência foi de 5,5 e não é truncado, ou seja, ocorreu Censura;
- O cliente *b* também é observado no início do estudo e seguido até ao fim do 7º ano sem ter verificado o evento de interesse; o tempo de sobrevivência é truncado, há Censura, entretanto, apenas podemos dizer que não houve o evento de interesse no período observado;
- O cliente *c* é seguido até o 6º ano, o tempo de sobrevivência é truncado, ou seja, censurado em 6 anos;
- O cliente *d* é seguido até ao fim do estudo sem verificar o evento de interesse, portanto, também é censurado em 5 anos;
- O cliente *e* é seguido até ao fim do estudo sem verificar o evento de interesse, portanto, também é censurado em 5 anos;
- O cliente *f* é seguido até o 7º ano em que se observa o evento de interesse, portanto seu tempo de sobrevivência é 7 anos no estudo.

Desse modo podemos observar pelo esquema demonstrado que todos os clientes possuem início no mesmo tempo no estudo e, portanto assumindo o mesmo período e risco inicial. E assim podendo definir a partir dos dados censurados de cada indivíduo, os diferentes tempos de sobrevivência.

3.2.2 Censura

Censuras, conforme Colosimo e Giolo (2006) podem ocorrer antes que o evento de interesse aconteça em relação aos casos da amostra, são observações incompletas ou parciais do tempo de sobrevivência.

Para Kachman (1999), a Censura é a principal característica para modelagem de sobrevivência, pois define o tempo decorrido até um evento de interesse, e que tal fato representa alguns desafios, como por exemplo: dados com distribuições não normais e várias formas de censura.

Ainda Colosimo e Giolo (2006) informam que tais observações podem ocorrer por diversas razões e, em função disso, não ocorrer o evento de interesse, neste estudo, considerados inadimplência até a data final do estudo, ou seja, 70 meses de observações.

Desse modo conclui que as observações censuradas, possuem tempo de falha superior ou não do tempo estipulado do estudo. Entretanto o autor, afirma que todas as observações em um estudo de sobrevivência devem ser utilizadas e analisadas, uma vez que produzirão argumentos para as fórmulas matemáticas dos modelos de previsões e assim diferencia-se esta técnica das outras utilizadas em modelagem de risco de crédito, por exemplo, Regressão Logística.

Para Kleinbaum e Klein (1996), na maioria das análises de sobrevivência deve-se considerar que a problemática está centrada nas condições observadas de Censuras, e que ocorrem devido a alguma informação sobre o tempo de sobrevivência individual, mas não se sabe o tempo de sobrevivência exatamente. Segundo o autor há três tipos conhecidos de Censura: tipo I, denominada censura à direita, quando o estudo é terminado após um período

preestabelecido de tempo, a tipo II é aquela onde o estudo será terminado após ter ocorrido o evento de interesse, e do tipo aleatória, tipo III, que é o mecanismo de Censura mais comum na Medicina ou Engenharia, ou seja, a observação é retirada no decorrer do estudo sem ter ocorrido o evento de interesse. Neste estudo a censura observada é a do tipo I.

3.2.3 Funções do tempo de Sobrevivência e Risco

Segundo Everitt e Hothorn (2010), a principal variável do resultado observado é o tempo para a ocorrência de um evento específico, é a importância central nos estudos de análises de sobrevivência sendo representada por duas funções: Função de Sobrevivência e Função de Risco, ambas utilizadas para descrever a distribuição dos dados, sendo denominada função de sobrevivência.

O tempo de sobrevivência de um indivíduo (neste estudo é o tempo de adimplência de um usuário de cartão de crédito) é uma variável aleatória T que pode assumir a função de valores não negativos, conforme explica Colosimo e Giolo (2006).

Estes valores T podem assumir uma distribuição de probabilidade que pode ser especificada de várias formas, duas das quais são particularmente úteis e bastante usadas para ilustrar diferentes aspectos dos dados em aplicações de sobrevivência, conforme já mencionadas por Everitt e Hothorn (2010) a função de sobrevivência e a função de risco, também descritas no modelo matemático de Bellotti e Crook (2007), (ver 3.4 Modelo matemático).

Para Kachman (1999) é por meio da função de risco que se mede o risco de ocorrer o evento no tempo t , a inadimplência, sendo que a probabilidade de falha até o tempo t é dada pela função de sobrevivência no modelo.

3.2.4 Estimação da função de sobrevivência

Estudos onde a estimação de tempo de sobrevivência é fator primordial, frequentemente apresentam-se observações censuradas, e, portanto são necessárias técnicas estatísticas específicas para analisar dados censurados nas observações. A estimação deste tempo é frequentemente analisada em estudos clínicos, o estimador mais utilizado é o proposto por

Kaplan-Meier (1958) conforme descrito em Colosimo e Giolo (2006). Esse estimador é conhecido como não paramétrico e também chamado de estimador limite-produto, pois na sua construção considera tantos intervalos de tempo quantos forem o número de falhas distintas, sendo que os limites dos intervalos de tempo é considerado como os tempos de falha da amostra.

3.2.5 Modelo de riscos proporcionais de Cox

Segundo Kleinbaum e Klein (1996), a principal razão pela qual o modelo de Cox é popular e amplamente utilizado é por ser semiparamétrico e, portanto mesmo que a função base de risco não seja especificada, fornece estimativas razoavelmente boas dos coeficientes de regressão, das taxas de risco em relação ao evento de interesse, e por último poderá ajustar as curvas de sobrevivência em detrimento a uma grande variedade de dados e situações. Ao ser comparado com os outros modelos, conforme o autor, o modelo de risco proporcional de Cox é considerado "robusto", ao ser comparado aos resultados de modelos paramétricos.

Portanto a expectativa principal na utilização destes modelos é que podem ser adaptados de forma a permitir influência de variáveis explicativas como covariáveis e deste modo aumentar o poder preditivo dos modelos de risco, portanto é considerado como modelo de risco proporcional de Cox.

Segundo Kleinbaum e Klein (1996), essas covariáveis devem ser incluídas na análise estatística dos dados com a finalidade exclusiva de explicação do possível efeito no tempo de sobrevivência.

O modelo de regressão de Cox é formado pelos coeficientes β que é o vetor de parâmetros associados às covariáveis, e, portanto garante o efeito das covariáveis sobre a função de risco no modelo, entretanto em relação ao método de estimação, que não se pode realizar através do método de máxima verossimilhança, devido à presença do componente não paramétrico é necessário utilizar a função de verossimilhança parcial.

A função de verossimilhança "parcial", ou seja, probabilidade "parcial" é usada por se considerar somente às probabilidades para os indivíduos que não falharam ao longo do período observado e, portanto não são considerados os indivíduos censurados. Assim,

a probabilidade para o modelo de Cox não considera probabilidades para todos os indivíduos, portanto, é chamado de probabilidade parcial (Kleinbaum e Klein, 1996, p.99).

3.2.6 Covariáveis dependentes do tempo

Colosimo e Giolo (2006) descrevem que ao utilizar variáveis em um determinado estudo, e que tais variáveis possam mudar ao longo do tempo, denomina-se como covariáveis dependentes do tempo. O objetivo da inclusão destas covariáveis no estudo é pela possibilidade de modelar seu efeito ao longo do período observado, e, portanto, fornecer estimativas da sensibilidade ou não ao modelo.

Neste estudo o tempo de sobrevivência será o início do período observado na amostra até o primeiro incumprimento, neste caso, o evento de interesse, a inadimplência por 60 dias de atraso no pagamento.

As variáveis dependentes do tempo são utilizadas para acessar a suposição de risco proporcional na função de covariável dependente do tempo, o modelo de Cox então, é estendido com o propósito de conter esta interação com o intuito de observar seu comportamento na função do tempo.

Por exemplo, se a variável Sexo - é utilizada como variável somente como estratificação dos dados em relação à suposição de risco proporcional - o modelo Cox deverá ser estendido para conter a variável Sexo $\times t$, adicionando a variável Sexo. Se o coeficiente da covariável analisada se tornar significante, podemos concluir que a suposição de risco proporcional foi violada pela covariável Sexo. (Kleinbaum e Klein, 1996, p. 136).

3.2.7 Modelo de Cox Estratificado

A suposição inicial para utilizar o modelo de Cox Estratificado, Colosimo e Giolo (2006) é que a suposição de riscos proporcionais é violada, ou seja, assumir que a amostra está dividida em estratos e que há um risco basal específico para cada estrato, por exemplo: as taxas de falha podem não ser proporcionais entre os homens e mulheres.

Entretanto esta suposição pode valer-nos estratos formados somente por homens ou por mulheres conforme visto em (Colosimo e Giolo, 2006).

Outra importante suposição é assumir que os efeitos das covariáveis não variam entre os estratos, ou seja, as covariáveis atuam de maneira similar na função de taxa de falha de cada estrato. Ainda Kleinbaum e Klein (1996) informam que será utilizada a função de risco para estimar os parâmetros do modelo.

3.3 Base de dados

Neste estudo será considerada uma amostra de dados brasileiros que compreendem uma população de indivíduos que possui limite de crédito por meio da abertura de conta de cartão de crédito. As variáveis foram classificadas como dependentes e independentes, onde se referem ao enquadramento ou não do cliente como inadimplente e aos fatores que influenciaram estes resultados.

O estudo apresenta um total de 18 variáveis das quais, apenas 6 foram selecionadas de acordo com o objetivo proposto neste trabalho, conforme demonstradas abaixo:

Tabela 4: Variáveis da amostra

1	Ref.	Data do início do estudo
2	Cliente	Portador do cartão de crédito, considerado apenas o titular do cartão de crédito, identificado numericamente pela amostra.
3	Sexo	Feminino/Masculino.
4	Limite	Limite de Crédito. Valor limite a ser utilizado pelo usuário do cartão de crédito. Limite estabelecido pela Administradora no Contrato, ocasião da contratação e de acordo com as características de Cadastro apresentado pelo cliente em conjunto com as regras internas de concessão de crédito da Administradora.
5	Financ.	Indicador de utilização do crédito rotativo. Parcelamento do valor total da fatura,

		mínimo 15% do valor para pagamento. Entretanto tal situação não é considerada inadimplência em relação à utilização de cartão de crédito
6	Data	Data de Nascimento.
7	Prof.	Profissão.
8	Estado	Território brasileiro na abertura da conta de cartão de crédito.
9	Canal	Classificação referente à forma de venda do cartão de crédito pela instituição financeira no momento da contratação do crédito pelo cliente.
10	Renda Dcl.	Renda declarada no momento de abertura do cartão de crédito pelo cliente, sem apresentação de comprovante de Renda.
11	Renda Conf.	Renda Confirmada, apresentação de comprovante de renda pelo cliente.
12	Saldo	Saldo disponível em relação ao Limite de Crédito.
13	Normal	Conta ativa. Conta que não apresenta nenhuma restrição de crédito e utilização pelo cliente observado pela Administradora do cartão.
14	Cancel.	Conta cancelada. Cartão que por algum motivo foi solicitado o cancelamento do cartão de crédito pelo cliente.
15	Enquad.	Cliente Inadimplente. Conta “Enquadrada” nas condições de inadimplência estabelecidas pela Administradora na ocasião da contratação do cartão. (60 dias de atraso)
16	Taxa	Taxa de juros da Administradora de cartão de crédito.
17	Cartões	Qtd. de cartões de crédito contratados por cada cliente.
18	Abertura	Data da abertura da conta do cartão de crédito pelo cliente junto á Administradora.

Na Tabela 4 podemos observar que quase a totalidade de nossa amostra é composta por variáveis que descrevem o relacionamento do cliente com a instituição emissora do cartão de crédito.

As variáveis abaixo mencionadas não foram consideradas relevantes, de acordo com o objetivo proposto e, portanto foram descartadas da amostra, após análise detalhada dos dados:

- Data: Data de Nascimento, não aplicável a pesquisa;
- Profissão (Prof.): não aplicável à pesquisa;
- Estado: não aplicável à pesquisa;
- Saldo: não aplicável à pesquisa;
- Cancel (cancelamento): não considerada na pesquisa como variável relevante para inadimplência, entretanto, pode-se considerar indicadora de Censura na amostra.
- Cartões: cada cliente apresenta em sua maioria um único cartão de crédito
- Abertura: não aplicável. A data considerada para início deste estudo foi estipulada em 3 safras Maio/Junho/Julho de 2006, conforme dados disponíveis. A safra escolhida refere-se ao período de Maio de 2006 á Fevereiro de 2012.
- Financ: Financiamento, crédito rotativo. Variável descartada por não configurar inadimplência em cartões de crédito;
- Canal: venda realizada pela gerencia, canal 4, grande maioria dos clientes analisados;
- Renda Decl (Renda Declarada): *outliers*, informações incompletas.
- Taxa de Juros: Aplicada conforme relacionamento, cliente e administradora, em relação à data (período) no momento da inadimplência.

Na Tabela 5, são demonstradas as variáveis escolhidas, após análise da base de dados:

Tabela 5: Variáveis da pesquisa

1	Cliente	Portador do cartão de crédito, considerado apenas o titular do cartão de crédito, identificado numericamente pela amostra.
2	Tempo de Sobrevivência	Data de Início e Fim deste estudo, 01.05.2006 á 01.02.2012. (70 meses) Variável

		indicadora do Tempo de Sobrevivência de cada indivíduo na amostra
3	Sexo	Feminino/Masculino.
4	Limite	Limite de Crédito. Valor limite a ser utilizado pelo usuário do cartão de crédito. Limite estabelecido pela Administradora no Contrato, na ocasião da contratação e de acordo com as características de cadastro e regras internas de concessão de crédito da Administradora.
5	Enquadramento	Inadimplente. Variável indicadora de Inadimplência na amostra. Considerando que o cliente está em atraso no pagamento da fatura do cartão de crédito por mais de 60 dias de atraso.
6	Renda Confirmada	Renda Confirmada, apresentação de comprovante de renda pelo cliente no momento da contratação do cartão de crédito.

A justificativa da escolha das variáveis, conforme descrito anteriormente, foi resultado da análise dos dados disponíveis, a adequação da proposta desta pesquisa, prática no mercado bancário de comprovação de Renda e Fundamentação Teórica, conforme demonstrado na Tabela 2.

3.3.1 Variáveis macroeconômicas

As variáveis macroeconômicas dependentes do tempo incluídas na análise são: taxa de juro real da economia, taxa de desemprego e taxa de inflação medida pelo IPCA. Para manter a compatibilidade com a amostra de dados de cartão de crédito, os dados são mensais no período de maio de 2006 a fevereiro de 2012. As séries macroeconômicas foram obtidas no Instituto Brasileiro de Geografia e Estatística (IBGE), Ipeadata e Banco Central do Brasil.

Tabela 6: Variáveis macroeconômicas

Qtd.	Variável macroeconômica	Fonte	Formato
01	Taxa de Juros Selic	Banco Central	Mensal
02	IPCA - Índice Nacional de Preços ao Consumidor - Amplo	Banco Central	Mensal
03	Taxa de Desemprego	IBGE	Mensal

Ao incorporarem a taxa de juros, nível de desemprego dentre outras variáveis macroeconômicas, em modelos de Risco de Crédito em Análise de Sobrevivência, os autores procuram observar se tais variáveis escolhidas podem ser consideradas fatores de influência positiva ou negativa em relação ao risco de *default*, neste caso, inadimplência, cuja definição segundo Annibal (2009), é complexa e menciona que o mais indicado para medir inadimplência no *stricto sensu* seria identificar, em relação ao segmento bancário, seria o número de operações em atraso.

Com o aumento na confiança do consumidor em geral, espera-se conduzir a um maior risco relativo, pois os consumidores ficam mais propensos a consumir e pode ser difícil realizar os pagamentos, conforme observado por Tang *et al.* (2006) e Rosch e Scheule (2004).

Porém ainda em Rosch e Scheule (2004), a utilização de variáveis macroeconômicas serve para demonstrar o efeito esperado ou a expectativa de efeito esperado sobre a probabilidade de um risco de *default* pelo consumidor tomador de crédito, neste caso, cartões de crédito. A justificativa para o uso de variáveis macroeconômicas se dá pelo fato da possibilidade de uma determinada variável macroeconômica influenciar no risco de default pelo tomador de crédito.

Desse modo Thomas e Stepanova (2002) afirmam que se espera que as taxas de juros possam ter maior influência no risco, uma vez que esta variável pode ter efeito decisivo em relação a pagamentos de dívidas.

Segundo Thomas (2001; 2005), espera-se que as variáveis: alta taxa de juros de um país, alto nível de desemprego, variação do índice imobiliário tenha importante impacto na possibilidade de um evento de *default*, traduzindo-se em alto risco de inadimplência.

Considera-se, porém, que o inverso também seja verdadeiro, caso as mesmas variáveis, por exemplo, alta de juros e nível de desemprego diminuam em um determinado período, estas também poderão influenciar no risco de inadimplência, e assim possibilitar a revisão de crédito pelo credor.

Por último conforme observa, Banasik *et al.* (1999) a intenção é demonstrar se a inclusão de variáveis macroeconômicas incorporadas a um modelo de sobrevivência como variáveis no tempo, covariáveis, fornecerá ao modelo, resultado estatisticamente significativo.

3.4 Modelo matemático

O modelo de regressão de Cox (Cox, 1972), modelo escolhido para este trabalho é utilizado para analisar dados censurados de sobrevivência conforme Gil (1984) e de acordo com Colosimo e Giolo (2006), que considera como Censura é a observação parcial do tempo de interesse de um conjunto de dados sendo a variável resposta o tempo.

O modelo é o semiparamétrico, conhecido como modelo de riscos proporcionais de Cox, e apresenta uma parte não paramétrica composta por uma constante e uma parte paramétrica, composta pelas variáveis explicativas, conforme os trabalhos de Whalen (1991) e Rocha (1999).

O modelo prescreve que o tempo de sobrevivência é medido a partir da data em que a conta foi aberta até a data de corte dos dados. Se uma conta não tiver evento de falha, ou seja, evento de interesse, durante o período da abertura da conta do cartão de crédito até a data de corte da amostra, estes dados serão considerados censurados.

No contexto de crédito ao consumo, os dados analisados são de indivíduos que obtêm crédito na forma de empréstimos por meio de cartões de crédito. Bellotti e Crook (2007) propuseram as seguintes equações: em dados de sobrevivência por meio da função de risco, a equação (1) fornece a variação da taxa de probabilidade de falha no tempo t :

(1)

$$h(t) = \lim_{\delta t \rightarrow 0} \left(\frac{P(t \leq T < t + \delta t \mid T \geq t)}{\delta t} \right)$$

Onde T é uma variável aleatória associada ao tempo de sobrevivência. A probabilidade de sobrevivência no tempo t pode ser dada em termos da função de risco na equação (2):

(2)

$$S(t) = P(T \geq t) = \exp\left(-\int_0^t h(u)du\right).$$

Segundo Bellotti e Crook (2007), a equação (2) permite observar a probabilidade de sobrevivência até ao tempo t , em relação aos dados de crédito dados. É, portanto, a probabilidade de que não ocorreu o evento de inadimplência por parte de alguns membros da população em um determinado período de tempo t , ou seja, após a abertura da conta de cartão de crédito até a qualquer momento, sendo 1- PD no tempo t , conforme Whalen (1991).

Em relação à equação (3), Bellotti e Crook (2007) utilizaram o modelo semiparamétrico de Cox, (COX, 1972) para possibilitar a inclusão de variáveis macroeconômicas como variáveis no tempo – covariáveis. O modelo depende parcialmente de um vetor de coeficientes β , que são múltiplos lineares das covariáveis, e também de uma função de risco h_0 , dependente do tempo, mas não das covariáveis:

(3)

$$h(t, \mathbf{x}(t), \beta) = h_0(t) \exp((\beta \cdot \mathbf{x}(t)))$$

A equação (3) fornece o risco no tempo t das observações \mathbf{x} , de acordo com os parâmetros de β . Bellotti e Crook (2007) mencionam que o modelo produz uma função estimativa de probabilidade (máxima verossimilhança), *Likelihood* (COX, 1972), parcial sobre as observações de teste na equação. No presente trabalho o vetor X corresponde a um conjunto

de índices macroeconômicos e t corresponde ao tempo de sobrevivência de um indivíduo da amostra.

$$l_p(\boldsymbol{\beta}) = \prod_{i=1}^n \left(\frac{\exp(\boldsymbol{\beta} \cdot \mathbf{X}_i(t_{(i)}))}{\sum_{j \in R(t_{(i)})} \exp(\boldsymbol{\beta} \cdot \mathbf{X}_j(t_{(i)}))} \right)^{c_i} \quad (4)$$

Desta forma $t_{(i)}$ são ordenados como tempos de sobrevivência e o risco como: $R(t) = \{j: t_{(j)} \geq t\}$, permitindo a máxima verossimilhança para estimar $\boldsymbol{\beta}$ sem a necessidade de conhecer a taxa de risco, porém, em função do modelo ser para estimar sobrevivência é necessário conhecê-la para estimar os parâmetros do $\hat{\boldsymbol{\beta}}$ do $\boldsymbol{\beta}$ pela estimação máxima da verossimilhança.

Bellotti e Crook (2007) mencionam que uma série de observações de n $i=1$ é dada na Análise de Sobrevivência em termos de quantidade de observação no tempo t_i e indicadores c_i onde $c_i = 0$ para uma observação censurada e $c_i = 1$ para um evento de falha, caso em que t_i é o tempo de sobrevida. Ainda mencionam que cada observação incluirá um vetor de covariáveis que podem estar associados ao tempo de sobrevivência, sendo também variáveis no tempo e em geral, conforme os autores são dados como funções do tempo, $x_i(t)$ e ainda as variáveis macroeconômicas podem mudar com o tempo e o valor da covariável é dado como o valor das variáveis macroeconômicas, no momento da falha.

Os indivíduos que não entraram em *default* durante o período são classificados como censurados e seu tempo de sobrevivência é o último período em que foram observados. A estimação dos parâmetros do modelo é dificultada pela presença do componente não paramétrico, e, portanto o modelo de Cox propõe um método de máxima verossimilhança parcial, que considera o conhecimento da história passada de falhas na construção da função de máxima verossimilhança, Colosimo e Giolo (2006).

Para este estudo, na aplicação do modelo com dados brasileiros, são utilizadas as variáveis macroeconômicas: Taxa de Desemprego, Taxa de Juros – Selic e IPCA, (Índice nacional de

Preços ao consumidor - Amplo) conforme estudos anteriores apresentados na Tabela 2, e porque são as variáveis consideradas as que mais se aproximam do objetivo proposto neste estudo, o risco de inadimplência.

3.4.1 Forma de análise dos resultados

Modelos de risco de crédito exploram hipóteses causais do comportamento de tomadores de crédito em relação à sua capacidade de pagamento. Instituições financeiras e agentes reguladores utilizam esses modelos para avaliar e estimar o risco individual ou possíveis perdas nas carteiras de crédito. Desta forma, os bancos podem acessar informações sobre possíveis perdas futuras e calcular sua provisão de crédito. Utilizando o modelo descrito, o interesse é avaliar se as variáveis macroeconômicas utilizadas no modelo possuem poder explicativo em relação à previsão de *default*.

Para estimar a probabilidade de *default* de cada cliente ao longo do período observado, foi utilizado o Modelo de Regressão Estratificado de Cox. O tempo de sobrevivência corresponde ao tempo de sobrevivência entre o período observado em que o cliente leva para tornar-se inadimplente desde a sua entrada no estudo.

Considerando as normas de mercado e normas internas da instituição emissora do cartão de crédito detentora da base de dados, foi definida a inadimplência a partir de 60 dias de atraso no pagamento da fatura. Os clientes que não atingiram esse nível de atraso no período analisado foram considerados censurados. Nesse estudo, o tempo de sobrevivência é definido como a quantidade de meses em que foram observados, tendo como limite: a) data do corte da amostra de 01 de maio de 2006 para a primeira safra, considerada no período final das observações da amostra.

As análises foram realizadas utilizando o Software R versão 2.4.1 que se caracteriza pelo compromisso entre a grande flexibilidade e a conveniência de *softwares* estatísticos tradicionais, incluindo ampla variedade de métodos estatísticos tradicionais e modernos. A versão autorizada foi a 2.4.1 e está disponível no endereço [HTTP://.www.r-project.org](http://www.r-project.org), para a análise das variáveis macroeconômicas foi utilizado o software estatístico Stata Corporation, versão,8.0.

4 RESULTADOS

Para aplicar as técnicas propostas neste estudo, foram utilizadas amostras de clientes correntistas portadores de cartão de crédito de uma instituição financeira. Essas amostras são compostas de dados mensais de clientes entre maio de 2006 a fevereiro de 2012, compreendendo um total de 3.396 clientes, divididos em três safras. A safra 1 é composta por clientes que contrataram o cartão de crédito no mês de maio/2006, a safra 2, por clientes que contrataram o cartão de crédito no mês de junho/2006 e a safra 3, com contratação do cartão no mês de julho/2006.

Os resultados discutidos nesse trabalho usam dados da safra 1 que compreendem um total de 1.418 indivíduos, sendo a safra com o maior número de clientes. Análise dos dados de outras safras não conduz a resultados distintos e, portanto, não são reportados. É importante salientar que safras diferentes poderiam gerar modelos e resultados substancialmente distintos caso fossem associadas a estratégias de venda diferentes como, por exemplo, promoções ou esforços diferenciados na colocação do cartão.

Conforme Tabela 5, as variáveis da amostra foram escolhidas conforme disponibilidade dos dados e referencial teórico, demonstrados também no Tabela 2.

Na Tabela 7 são demonstrados os dados da amostra compilados no formato de matriz a ser utilizado pelo software R para análise descritiva dos dados.

A variável Cliente é utilizada em forma numeral identificando cada indivíduo da amostra ao longo de 70 meses, prazo final das observações; a variável Sexo estratifica os indivíduos da amostra em feminino e masculino; a variável Tempo de Sobrev. (Sobrevivência) é a variável que indicadora do tempo de sobrevivência de cada indivíduo na amostra. Esta variável foi construída considerando o tempo inicial do estudo e a data do primeiro Evento de Interesse, de cada indivíduo na amostra.

A variável Enquad (Enquadramento) principal variável da amostra, pois indicará de ocorreu ou não o evento de interesse. Considera-se 0 para dados censurados e 1 para indicação da

ocorrência do evento de interesse, inadimplência; renda Conf. (Renda Confirmada) variável que indica o valor atribuído como Renda no momento da ocorrência do Evento de Interesse.

A variável Limite de Crédito indicará o montante atribuído pela Administradora de cartão de crédito em relação ao limite de crédito para consumo de cada indivíduo da amostra.

Os dados a partir da matriz serão manipulados através do software R, demonstrado na tabela 7:

Tabela 7: Matriz dos dados da amostra

Cliente	Sexo	Tempo de Sobrev.	Enquad	Renda Conf	Limite
1	F	70	0	10000	7000
2	M	70	0	1500	750
3	F	70	0	4000	2500
4	F	45	1	1850	900
5	F	47	1	3500	2400

Tempo de sobrevivência é o resultado esperado entre o início do tempo observado até o término do estudo, considerado observações mensais (30 dias). Neste estudo o tempo máximo de sobrevivência do cliente será de 70 meses.

A técnica de análise de sobrevivência permite avaliar clientes que sobrevivem mais do que o limite de tempo da pesquisa considerando que durante o tempo observado não ocorreu o evento de interesse, neste caso, a inadimplência. A Tabela 8 apresenta as estatísticas descritivas das variáveis utilizadas na amostra.

Tabela 8 : Análise descritiva dos dados da amostra

Tempo de Sobrev.	Enquadramento	Renda Conf.	Limite
Min. 3.00	Min.: 0.0	Min.: 350.00	Min.: 200
1st Qu.: 0.00	1st Qu.:0.00	1st Qu.:2.200	1st Qu.: 1500
Median: 7 0.00	Median:0.00	Median:3.000	Median : 2000
Mean: 54.19	Mean: 0.3543	Mean.: 3.844	Mean : 2539
3rd Qu.: 70.00	3rd Qu: 1.000	3rd Qu: 5.000	3rd Qu.: 3300
Max: 70.00	Max.: 1.000	Max.:20.000	Max. :11.700

Observa-se na Tabela 8 que o tempo mínimo de sobrevivência é 3 meses com uma média de aproximadamente 54 meses de sobrevivência sem que tenha ocorrido o evento de interesse.

O Enquadramento, ou seja, a variável indicadora de inadimplência foi considerada a data do primeiro Evento de Interesse no período observado, denominada na amostra como 0 ou 1, sendo 0 para dados censurados e 1 para dados enquadrados. Obteve, após testes descritivos estatísticos, uma média de 0.3543 de significância em relação à amostra de 1.418 clientes, ou seja, aproximadamente 35% dos clientes se tornaram inadimplentes durante o período observado.

Os dados analisados apresentaram variações nos valores atribuídos as variáveis, Limite e Renda Confirmada, portanto foram considerados aqueles que constavam em seus registros no momento da ocorrência do evento de interesse.

Clientes que possuíam registros com valores de limite de crédito e renda confirmada acima do valor estipulado na amostra, R\$25.000 foram descartados da amostra, tendo sido considerados *outliers*.

4.1 Funções de sobrevivência

Primeiramente definimos por T a variável aleatória significando o tempo de cliente sobreviver à inadimplência, T pode ser qualquer número igual ou maior do que zero. Em seguida, t é

intervalos de tempo quantos forem o número de falhas distintas, calculando a sobrevivência toda vez que há um evento de interesse.

Tabela 10: Estimando função de sobrevivência

Time	n.risk	n.event	Survival	Std.err	Lower 95% CI	Upper 95% CI
3.00	1418	3	0.998	0.00121	0.995	1.000
3.07	1415	18	0.985	0.00321	0.979	0.991
4.00	1397	6	0.981	0.00363	0.974	0.988
4.10	1391	8	0.975	0.00412	0.967	0.983
5.00	1393	5	0.972	0.00437	0.963	0.981
5.10	1383	6	0.961	0.00445	0.962	0.980
5.10	1377	10	0.964	0.00494	0.954	0.974

A Tabela 10 demonstra que a primeira coluna (*Time*) indica o tempo em que diferentes indivíduos estão em seguimento no estudo, ordenados por tempos crescentes. A segunda coluna (*n.risk*) indica os indivíduos restantes no estudo em risco ou não (Censurado).

A coluna seguinte indicadora dos eventos de interesse (*n.event*) (*Cumulative events*) contém os eventos acumulados, isto é, os que se tornaram inadimplentes até este tempo, a coluna (*survival- Cumulative survival*) - Sobrevivência acumulada é a sobrevivência acumulada ou a proporção de casos para os quais não ocorreu o evento em cada tempo.

A coluna (*Standard Error*) é o erro padrão correspondente à estimação pontual de Kaplan-Meier em cada tempo e por último são dados referentes às estimativas dos intervalos de confiança a 0.95 (*95% Confidence Interval*).

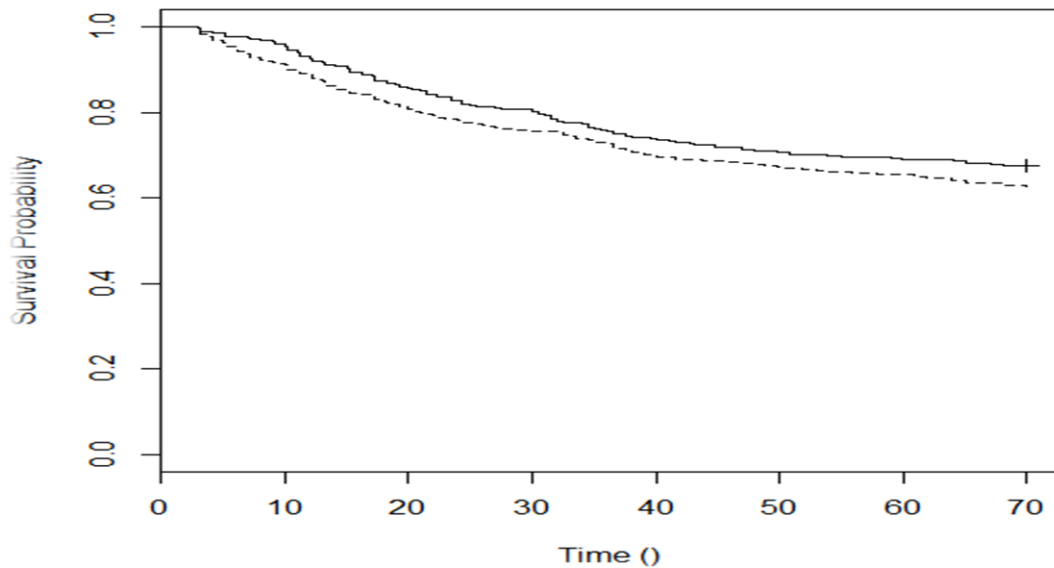


Figura 1: Curvas de Kaplan-Meier, sobrevivência por meses.

As funções de sobrevivência demonstradas na Tabela 10 e visualização gráfica, Figura 1, fornece a possibilidade de estimação da probabilidade de sobrevivência acumulada no período observado, bem como a curva de sobrevivência destas probabilidades *versus* o tempo de sobrevivência.

É a representação gráfica da função de sobrevivência $S(t)$ no eixo vertical *versus*, os tempos de sobrevivência (t), em meses, no eixo horizontal.

4.2.1. Estimativas de sobrevivência de Kaplan-Meier por fatores

Para que possamos comparar as Curvas de Kaplan – Meier e testar a significância estatística da relação de igualdade entre as curvas de sobrevivência a amostra será estratificada pelo fator Sexo. Nesta análise devemos dividir o conjunto total de observações em grupos distintos, sendo eles: grupo feminino e grupo masculino. Em seguida as funções de sobrevivência são estimadas separadamente para cada grupo e determinadas pelas curvas de sobrevivência de Kaplan-Meier.

Para testar a significância desta análise, será utilizado o teste de *Logrank*, conhecido como o teste mais utilizado para analisar a significância estatística em relação à Curva de Kaplan-Meier. O *Logrank test*, também conhecido por estatística do *logrank de Mantel*, estatística do *logrank de Cox-Mantel*, conforme Kleinbaum e Klein (2006).

A amostra é composta por 557 clientes do sexo feminino e 861 do sexo masculino e ao serem estratificadas pelos grupos feminino e masculino, o teste de *Log-Rank* irá comparar o número de eventos observados em cada grupo com o número de eventos esperados ordenando os tempos de "falhas" dos indivíduos nos dois grupos.

4.2.2 Teste estatístico de *Log-Rank*

Conforme demonstrado na Figura 2, o número esperado de falhas é calculado para cada intervalo para cada grupo, calculado em χ^2 entre as falhas esperadas *versus* falhas observadas e por fim assume o intervalo de tempo, conforme explicado em Kleinbaum e Klein, (2006, p.61).

Geralmente ao se aplicar testes de significâncias define-se a hipótese nula, neste estudo será a hipótese de que os grupos comparados possuem uma sobrevivência igual. O teste de hipótese *logrank* levará em consideração as diferenças de sobrevivência entre os grupos em todos os pontos em que dura à observação dos indivíduos na amostra.

O objetivo é responder: a sobrevivência foi diferente entre os fatores feminino e masculino? Ou há diferença significativa na sobrevivência entre homens e mulheres em relação à inadimplência em uso do cartão de crédito? E, portanto, define-se a hipótese nula.

H0: não há diferença entre as curvas.

H1: há diferença entre as curvas.

A curva de sobrevivência pode ser usada para comparar distribuições de sobrevivência dos grupos e fornece a possibilidade de determinar quantidades relevantes, por exemplo, mediana, *percentis* e outras medidas.

Através das curvas de Kaplan-Meier pode-se estimar diretamente o percentual de evento de interesse ao final de qualquer intervalo no tempo.

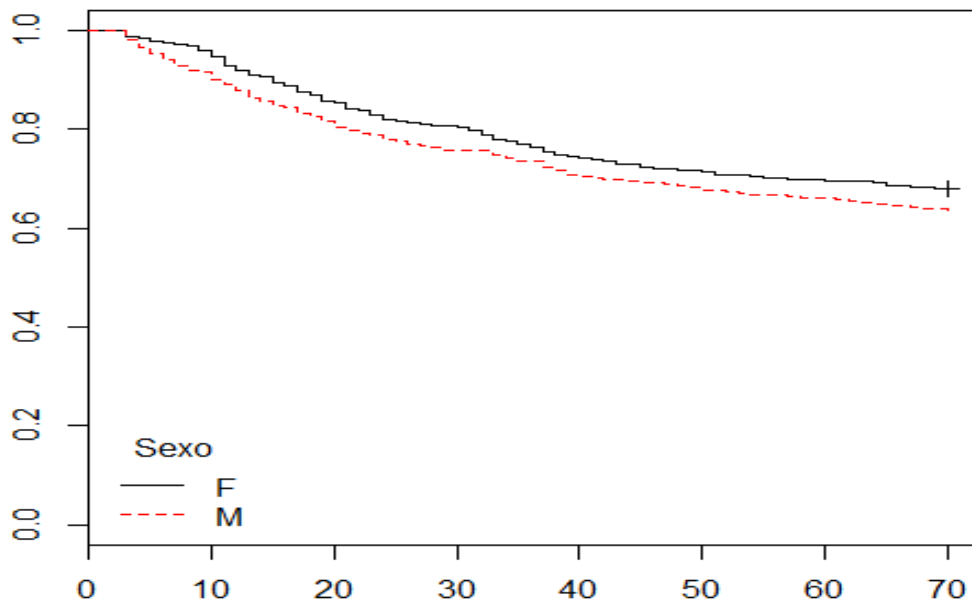


Figura 2: Kaplan Meier - estratificado por sexo

Podemos observar na Figura 3 que a curva de sobrevivência estratificada por sexo feminino e masculino no eixo vertical determina a probabilidade de sobrevivência, em porcentagem e no eixo horizontal determina a sobrevivência no tempo, mensal, período de 30 dias. O estimador Kaplan-Meier permite avaliar o ritmo em que ocorrem os eventos de interesse nos diferentes grupos em estudo.

Ao aplicarmos o teste *LogRank*, considerando o $\rho = 1$ com 1 grau de liberdade, sendo a probabilidade de não cometer o Erro do tipo I, ou seja, rejeitar a hipótese nula, as estatísticas extraídas são:

Tabela 11: Resultados Teste *Logrank*

Sexo	Qtd	N Observed	Expected	(O-E)^2/E	(O-)^2/V
F	557	181	203	2.34	3.95
M	861	321	299	1.59	3.95
Chisq= 3.9 on 1 degrees of freedom, p= 0.0469					

O resultado do teste *logrank* segue uma distribuição do Qui quadrado com 1 grau de liberdade, ao comparar as duas curvas na globalidade resultaram em um valor – $P = 0.0469$, permitindo, portanto rejeitar a hipótese nula. Assim, podemos concluir que há diferença significativa entre os sexos. Portanto permitindo afirmar que o sexo feminino apresenta maior sobrevivência em relação à inadimplência conforme testes e amostra analisada.

Entretanto tais resultados podem sugerir cautela ao considerar somente os resultados de significância estatística da amostra. Por exemplo, pode-se considerar que o comportamento do indivíduo em relação ao padrão de consumo é fator importante em relação à significância econômica dos resultados.

Por exemplo: a) indivíduos solteiros ou casados, b) utilização do cartão de crédito pelo membro que possui maior renda familiar, conforme IBGE (2012) as mulheres possuem Renda/Salário aproximadamente 28% menor do que os homens no mercado profissional. Nesse sentido considera-se o fato de que a mulher poderá utilizar o cartão de crédito do cônjuge para efetuar compras de maior valor, c) empréstimos do nome, ou seja, empréstimo do cartão de crédito á terceiros. Tais exemplos sugerem possíveis pesquisas e investigações sobre o tema.

4.3 Modelos de Regressão Paramétricos

Os modelos de regressão paramétricos aplicam-se a uma classe de modelos na análise de sobrevivência, denominados modelos paramétricos, em que o resultado segue uma distribuição normal, binomial, ou a distribuição de Poisson, ou seja, pertence a uma família distribuições de formas semelhantes com parâmetros desconhecidos. E somente quando o valor do parâmetro (s) é conhecido que se especifica totalmente a distribuição. Muitos desses

modelos paramétricos são modelos de aceleração do tempo de falha, que fornecem uma alternativa a estimar a taxa de risco sendo considerado “fator de aceleração”, conforme Kleinbaum e Klein (2006, p.258).

O objetivo principal da aplicação do modelo de regressão paramétrico em dados de sobrevivência é explorar a relação entre as covariáveis que podem estar relacionadas com o tempo de sobrevivência, considerada como variável resposta, que é o tempo de ocorrência de um evento de interesse, conforme Colosimo e Giolo (2006, p.115).

O pressuposto fundamental para Kleinbaum e Klein (2006), ao considerar um modelo de risco proporcional (PH), é que o risco é constante ao longo do tempo, e em contrapartida nos modelos paramétricos o pressuposto é que o tempo de sobrevida é acelerado ou desacelerado por um fator constante, quando comparados a diferentes covariáveis e por isso a distinção entre os modelos.

Na amostra estudada para a aplicação do modelo paramétrico, contém 1.418 clientes, sendo que o evento de interesse ocorreu para 502 (35.40%) dos casos e, conseqüentemente, 916 (65.02%) dos indivíduos foram censurados, clientes que não se tornaram inadimplentes no período estudado.

As variáveis escolhidas, dentre as variáveis potencialmente importantes, conforme demonstrado na Tabela 5, respectivamente, foram: Limite e Renda Confirmada.

Nesta etapa deste estudo o objetivo é analisar o efeito das variáveis (cadastro) no modelo de regressão paramétrico. Ou seja, verificar se as variáveis independentes (x) limite e renda confirmada são importantes para explicar a variável Y inadimplência. Sendo a hipótese nula considerando intervalo de confiança de 95%:

H0: Limite e Renda Confirmada não são importantes para explicar Inadimplência

H1: Limite e Renda Confirmada são importantes para explicar Inadimplência

Tabela 12: Resultados do modelo de Regressão Paramétrico

Variáveis	<i>t</i> - Função de Risco Haz.Ratio	Erro Padrão	z	p> z	95% I.C.	
Limite	,99981	0000421	-4,51	0,000	,9997275	,9998926
Renda Conf.	1,000102	,0000334	3,06	0,002	1,000037	1,000168
ln_p	-,0778142	,0423257	-1,84	0,066	-,1607711	,0051427
p	,9251363	,0391571			,851487	1,005156
1/p	1,080922	,0457508			,9948705	1,174416

Conforme resultados demonstrados na Tabela 11 e de acordo com os valores - *p* apresentados é possível rejeitar a hipótese nula, demonstrando, portanto que as variáveis são importantes para explicar inadimplência.

Entretanto em relação à função de risco, dado pelos coeficientes (*t-Haz.Ratio*) podemos observar que os resultados foram pouco diferentes de 1, permitindo interpretar que não houve aumento relevante em relação ao risco de inadimplência.

Porém a variável, limite apresentou um coeficiente muito baixo, (negativo) demonstrando possivelmente uma relação inversa no modelo permitindo interpretar que a variável possa diminuir o risco de inadimplência. Desse modo podemos interpretar que quanto maior o limite de crédito, menor o risco de inadimplência.

Por outro lado, tal resultado sugere outra possível interpretação, por exemplo, caso haja aumento no limite de crédito, esta fato poderá não significar menor risco de inadimplência, caso não ocorra mudança no padrão de consumo do indivíduo, haja vista, as premissas econômicas e estudos Keynesianos que argumentam que quando a Renda aumenta o Consumo também aumenta.

E por último, caso o cliente não consiga honrar seus pagamentos, este pode vir a solicitar o pagamento no crédito rotativo, que não configura inadimplência, mas invalida parcialmente o resultado de menor risco á inadimplência.

Assim, em relação ao modelo paramétrico acredita-se que as variáveis da amostra usualmente utilizadas como variáveis de cadastro não apresentam explicação para o risco de inadimplência em relação ao usuário de cartão de crédito, sendo, portanto passíveis a futuras investigações.

4.4 Teste de proporcionalidade do Modelo de Cox - Log-Log

Nesta etapa do estudo é necessário através de estimativas de curvas de sobrevivência ou de combinações de variáveis em uma determinada amostra, avaliar o pressuposto de proporcionalidade do modelo de risco a ser testado. O objetivo é afirmar se o modelo de risco proporcional é apropriado para a inclusão das variáveis, neste caso o limite de crédito e a renda confirmada como variáveis que não mudam ao longo do tempo, portanto estabelecendo uma dependência do tempo.

Para que possamos avaliar se o pressuposto de proporcionalidade do modelo de risco proporcional de Cox é atendido, é necessária à aplicação do teste de probabilidade de sobrevivência Log-Log (em conjunto com a análise dos resíduos, onde se verificará se há o paralelismo das curvas de sobrevivência indicando a função de risco de diferentes indivíduos em uma amostra).

Em Análise de Sobrevivência a probabilidade de sobreviver até qualquer ponto no tempo é estimada a partir da probabilidade cumulativa de sobreviver a cada um dos intervalos de tempo que precedem o evento de interesse, conforme explica Kleinbaum e Klein (2006, p.142). Entretanto o autor menciona que o paralelismo das funções de Log-Log de sobrevivência fornecerá uma abordagem gráfica para avaliar se o pressuposto de proporcionalidade do modelo de Cox (PH) é satisfeito, mas afirma que alguns problemas podem ocorrer em analisar a suposição de riscos proporcionais somente pela análise gráfica, sendo o principal problema, descobrir o quão paralelo às curvas estão.

Esta decisão conforme Kleinbaum e Klein (2006) poderá ser subjetiva para uma determinada base de dados, principalmente caso o tamanho da amostra seja pequena. Portanto

recomendam-se decisões cautelosas a respeito da não suposição, a não ser que haja fortes evidências de não paralelismo. A seguir o gráfico apresenta:

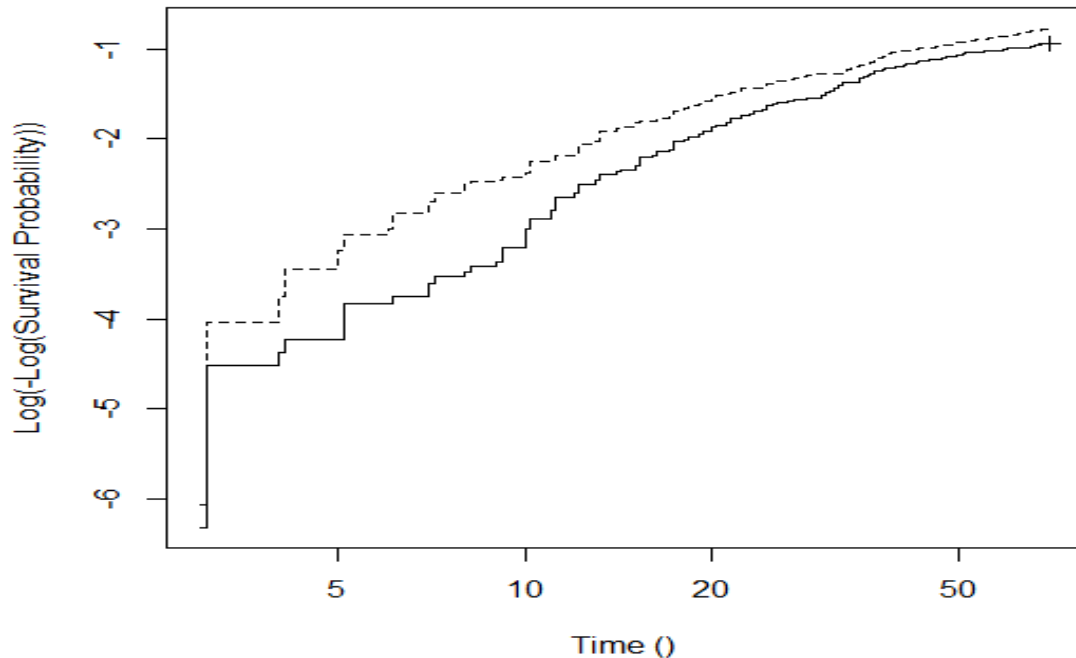


Figura 3: Probabilidade de sobrevivência Log-Log

As curvas de sobrevivência paralelas indicarão se o pressuposto de Risco Proporcional (PH) é satisfeito, ou seja, a soma linear das diferenças correspondentes aos valores preditivos para os dois indivíduos, sendo a distância vertical entre as duas curvas constantes e paralelas.

Entretanto, na Figura 4 não podemos afirmar que há um total paralelismo entre as curvas, representando os estratos feminino e masculino, e, portanto não podemos afirmar se o modelo de risco proporcional de Cox poderá assumir que a função de risco, ao comparar os dois estratos de preditores, é constante ao longo do tempo.

Dessa forma impossibilitando-nos de afirmar que o risco de um indivíduo é proporcional ao risco de qualquer outro da amostra, independentemente do tempo, sendo assim, realiza-se o teste Schoenfeld para dirimir esta dúvida.

4.4.1 Teste Schoenfeld - Análise dos Resíduos

O teste Schoenfeld analisará se o pressuposto de proporcionalidade de risco proporcional (PH) é satisfeito, o risco é comum para os dois estratos da amostra ao longo das observações.

Colosimo e Giolo (2006) mencionam que modelos incluindo covariáveis dependentes do tempo, não podem ser ajustados da mesma maneira como aqueles que incluem somente covariáveis que não mudam com o tempo, uma vez que tais covariáveis podem assumir diferentes valores em diferentes tempos e impactando no cálculo da função de verossimilhança parcial proposta no modelo de Cox.

Portanto para confirmar com segurança a análise gráfica, demonstrada na Figura 4, em relação ao pressuposto de proporcionalidade no modelo de risco proporcional de Cox, aplica-se o ajuste através do teste de resíduos padronizados Schoenfeld, conforme demonstrado a seguir:

Tabela 13: Teste de resíduos Schoenfeld

	rho	chisq	p
Limite	0,22840	34,58 1-df	0.000
Renda Conf.	-0,16425	16,38 1 df	0.0001
Global	NA	38,70 2 df	0.000

A ideia principal da suposição de risco proporcional é conter uma covariável que não esteja relacionada ao tempo de sobrevivência. E, portanto o teste de correlação dos resíduos é uma maneira de testar esta proporcionalidade analisando os resíduos da amostra.

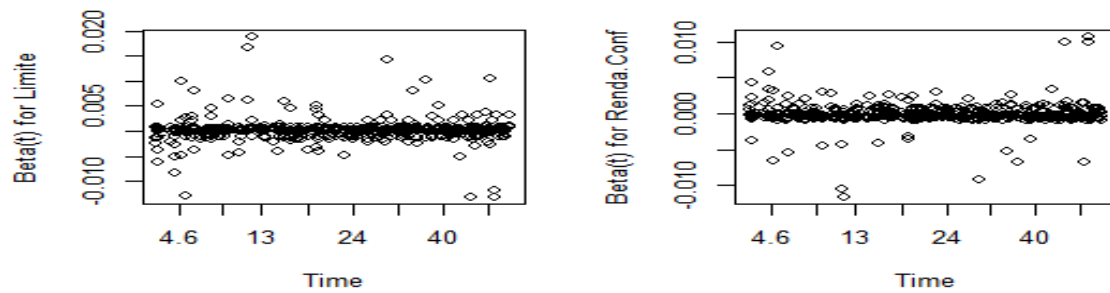


Figura 4: Teste de Resíduos de Schoenfeld do modelo de Cox com interação de limite e renda confirmada.

Define-se a hipótese nula do teste de proporcionalidade para o modelo de risco proporcional de Cox:

H0: O pressuposto de risco proporcional não é violado

H1: O pressuposto de risco proporcional é violado

Os resultados apresentados na Tabela 14 de acordo com os valores $-p$ observados apresentam evidências para rejeitar a hipótese nula de taxas de falha proporcionais, considerando intervalo de confiança de 95%.

Kleinbaum e Klein (2006, p.142) mencionam que a hipótese nula nunca é provada somente com testes estatísticos e que o mais comum é aceitar que não há fortes evidências para rejeitar a hipótese nula considerando que o resultado do valor $-p$ pode ser influenciado pelo tamanho da amostra. Desse modo, o autor sugere a análise estatística em conjunto com a gráfica para providenciar ao pesquisador bom senso em suas análises.

Portanto de acordo com os valores p apresentados permite-se rejeitar a hipótese nula de taxas de falha proporcionais considerando que houve violação no pressuposto de risco proporcional.

Desse modo conclui-se que o modelo de Cox com a interação das covariáveis não seria satisfatório para estes dados. Entretanto, conforme Kleinbaum e Klein (2006, p.142) sugere que o tamanho da amostra pode influenciar os resultados e que a visualização gráfica colabora para esta interpretação.

Decidiu-se, portanto ao analisar os resultados, considerando bom senso na decisão, e em conjunto com as imagens gráficas prosseguir com os estudos. O objetivo é a observação do comportamento das variáveis estudadas também no modelo de riscos proporcionais de Cox para futuras investigações.

4.4.2 Modelo semiparamétrico de risco proporcional de Cox

A suposição para o uso do modelo de regressão semiparamétrico de Cox é que a razão entre os grupos das covariáveis selecionadas no modelo (razão de risco) seja constantemente proporcional; no decorrer do estudo foi comprovada nos testes realizados nos itens 4.5 e 4.5.1 deste capítulo. Nesse caso, verificou-se a proporcionalidade de acordo com a função risco acumulada versus o tempo de cada covariável do modelo.

O principal interesse em estudos como este é avaliar como estas covariáveis influenciam o tempo de sobrevivência de cada indivíduo, e que possivelmente dependem da significância dessas covariáveis no modelo, agora sendo modeladas no modelo de riscos proporcionais semiparamétrico de Cox.

A seguir os dados descritivos estatísticos do modelo com as variáveis de cadastro, Limite de Crédito e Renda Confirmada:

Tabela 14 Resultados do modelo de risco proporcional de Cox – Limite e Renda Confirmada

Variáveis	t- Função de Risco (<i>Haz.Ratio</i>)	Erro Padrão	z	p > z	95% I.C.	
Limite	,9998928	,0000465	-2,30	0,021	,9998017	,999984
Renda Conf.	1,000045	,0000374	1,19	0,234	,9999712	1,000118

O objetivo nesta etapa do estudo é analisar o comportamento das variáveis da amostra no modelo semiparamétrico de Cox, verificar se a adição das covariáveis ao considerar o intervalo de confiança de 95%, irá interferir (aumentar ou diminuir) na função de risco do modelo, portanto estabelece a hipótese nula:

H0: Limite e Renda Confirmada não são importantes para explicar Inadimplência

H1: Limite e Renda Confirmada são importantes para explicar Inadimplência

Para Kleinbaum e Klein (2006, p.91), há basicamente três objetivos estatísticos a serem analisados em estudos como este, o primeiro é para testar a importância da variável escolhida como variável resposta, a segunda é a obtenção de uma estimativa do efeito dessa variável no evento de interesse analisado e a terceira é a determinação de um intervalo coerente para analisar tal efeito.

Os resultados demonstrados na Tabela 15 de acordo com os valores p apresentados em relação a variável Limite de Crédito permitiram rejeitar a hipótese nula, confirmando o resultado obtido no modelo paramétrico. Entretanto de acordo com os coeficientes analisados é possível também interpretar que a variável, Limite de Crédito estabeleceu uma relação inversa no modelo permitindo interpretar que esta variável diminui o risco de inadimplência, resultado também confirmado no modelo paramétrico.

A variável Renda Confirmada de acordo com os resultados apresentados não nos permitiu rejeitar a hipótese nula, portanto permitindo interpretar que a variável Renda Confirmada não é importante para explicar inadimplência, no modelo paramétrico. A variável apresentou

coeficientes, que atrelados à função de risco no modelo, demonstraram apenas 23% aproximadamente de poder de explicação em relação à inadimplência.

Esta interação pode ser considerada importante em relação a variável explicativa no modelo, uma vez que podemos interpretar que o indivíduo que apresenta Renda Confirmada aparentemente teria menor risco de inadimplência. Entretanto tal resultado também sugere questionamentos sobre a atual prática no mercado de crédito, principalmente bancário, onde a comprovação de Renda é obrigatória para concessão de crédito ao consumo ou financiamentos.

Em relação à função de risco no modelo, os coeficientes associados à taxa de risco apresentaram valores baixos, permitindo interpretar que devido aos coeficientes pouco diferentes de 1 não podem ser considerados significantes para explicar risco face à inadimplência.

4.4.3 Modelo de riscos proporcionais de Cox com variáveis macroeconômicas

Dado que o indivíduo sobreviveu até o evento de interesse, ou seja, até entrar em observação, nesta etapa, iremos analisar o comportamento das variáveis macroeconômicas incorporadas no modelo de risco proporcional de Cox, denominado estendido por Kleinbaum e Klein (2006). A forma geral do modelo PH Cox é que este modelo fornece a função de risco no tempo t para um indivíduo de acordo com as variáveis explicativas escolhidas e definidas como \mathbf{X} , ou vetor, preditor, sendo as variáveis modeladas para prever o risco individual de cada componente da amostra, ou seja, as variáveis interpretaram os efeitos de risco condicionado ao período de tempo.

O pressuposto é que as covariáveis podem ser monitoradas durante o estudo e seus valores podem mudar ao longo do tempo, ou seja, são dependentes do tempo e podem ser incorporadas ao modelo de regressão de Cox, desse modo, assumimos que o modelo não é mais de taxas de falhas proporcionais, sendo que a interpretação dos coeficientes β do modelo deve considerar o tempo t .

A intenção é demonstrar se a inclusão de variáveis macroeconômicas incorporadas a um modelo de sobrevivência fornecerá ao modelo resultado estatisticamente significativo. As variáveis macroeconômicas escolhidas foram: Taxa de Desemprego, Taxa de Juros - Selic e IPCA- Índice Nacional de Preços ao Consumidor – Amplo em conjunto com as variáveis da amostra, consideradas de cadastro, Limite de Crédito, Renda Confirmada e Sexo.

Tabela 15: Resultados do modelo de risco proporcional de Cox com variáveis macroeconômicas

Variáveis	t- Função de Risco Haz.Ratio	Erro Padrão	z p > z	95% I.C.	
Limite	,999859	,0000453	-3,11 0,002	,9997702	,9999478
Renda Conf.	1,000065	,0000363	1,80 0,072	,9999941	1,000137
Sexo	1,222338	,1138947	2,15 0,031	1,018306	1,467251
IPCA	1,301506	,3430013	1,00 0,317	,7764596	2,181591
Taxa de Juros -Selic	1,083567	,0436071	1,99 0,046	1,001382	1,172496
Desemprego	1,736148	,10689338	8,96 0,000	1,538787	1,958821
N= 1418 - Censuras: 502 Obs: 75685 Time at Risk: 75685 LR Chi2(6) = 118,72 Log Likelihood = -1483,8132 Prob > chi2 = 0,0000					

Conforme visto em Kleinbaum e Klein (2006, p.86) uma importante análise em estudos de sobrevivência é a verificação da Taxa de Risco ou Função de Risco. A taxa de risco é obtida pela função exponencial do coeficiente das variáveis para identificar o efeito das mesmas no modelo.

Portanto considera-se a hipótese nula, com o intervalo de confiança de 95%, de que não há alteração na função de risco com a inclusão das variáveis macroeconômicas, e que a função de risco, determinadas pelos coeficientes sejam diferente de 1.

A expectativa é que as variáveis macroeconômicas consideradas independentes “x” ao serem incorporadas no modelo de risco poderão explicar inadimplência, a variável dependente “Y”. De acordo com os resultados demonstrados na Tabela acima observamos que:

Em relação aos valores $-p$ dos testes indicam que é possível rejeitar a hipótese nula em relação às variáveis: Limite, Renda Confirmada, Sexo, Taxa de Juros-Selic e Desemprego, não permitindo, entretanto rejeitar para a variável IPCA.

Porém Colosimo e Giolo (2006 p.193) argumentam que há uma possível interpretação para não significância de algumas variáveis quando modeladas em conjunto, haja vista a importância suposta da covariável de acordo com a literatura existente, é que a covariável perca seu valor prognóstico na presença de outras covariáveis fortemente associadas. Portanto novos testes deverão ser realizados para confirmar tal argumento aplicado as variáveis separadamente. Neste estudo não foi realizado por não estar no objetivo principal do estudo.

E ao associarmos a função de risco, quase todas as variáveis incluídas no modelo apresentaram um coeficiente positivo, exceto a variável, Limite de Crédito, conforme visto nos modelos paramétricos e semiparamétrico de Cox. E em relação ao aumento no risco de inadimplência em função da inclusão das variáveis macroeconômicas no modelo, obtemos o seguinte resultado:

- Taxa de Desemprego - apresentou um aumento na taxa de risco de 1,73 vezes, portanto demonstrando ser a variável mais importante para explicar inadimplência;
- IPCA - apresentou um aumento na taxa de risco de 1,30 vezes, sendo considerada a segunda variável mais importante para explicar inadimplência;
- Taxa de Juros - Selic – apresentou um aumento na taxa de risco de 1,08 vezes demonstrando menor poder de explicação em relação à inadimplência considerando às demais variáveis macroeconômicas da amostra;
- Sexo - apresentou um aumento na taxa de risco de 1,2 vezes demonstrando baixa explicação de inadimplência em relação às variáveis macroeconômicas IPCA e Taxa de Desemprego consideradas as mais importantes do modelo.

As variáveis da amostra, denominadas Cadastro, (analisadas no modelo semiparamétrico de Cox realizado, sem a inclusão das variáveis macroeconômicas no item, 4.4.2), ao serem analisadas em conjunto com as variáveis macroeconômicas demonstraram pouca alteração em relação ao aumento de risco em relação à inadimplência. (ver item, 4.4.3).

5 CONSIDERAÇÕES FINAIS

Considerando o objetivo desta pesquisa de estudar a possibilidade de *default* em carteiras de crédito ao consumo através da técnica estatística de Análise de Sobrevivência verificou-se após testes e análises estatísticas que a técnica de Análise de Sobrevivência é adequada para investigação de dados a respeito de inadimplência. Os resultados demonstraram ser possível estimar com antecedência o momento de *default*, bem como fornecer classificação de risco em relação à previsão futura de risco de crédito individual devido à possibilidade de inclusão de variáveis macroeconômicas no modelo analisado.

A técnica de Análise de Sobrevivência se mostra muito útil na construção de modelos que levem em consideração a dimensão tempo em relação ao risco incorrido de um determinado evento, neste estudo, inadimplência. A técnica também se mostrou necessária para o desenvolvimento de modelos de classificação de risco de crédito principalmente aplicado à área financeira onde se percebe uma crescente expansão e demanda de crédito ao consumo e, portanto, a previsão de risco individual em carteiras de crédito seria um diferencial. Principalmente no que se refere à previsão do tempo médio de sobrevivência de uma carteira de crédito, conforme estudos de Andreeva (2006) e Thomas e Stepanova (2002) que também utilizaram a técnica de Análise de Sobrevivência no contexto financeiro e obtiveram resultados semelhantes.

Observou-se, entretanto que o tempo médio de sobrevivência dos indivíduos da amostra, usuários de cartões de crédito é de 54 meses e o mínimo de 3 meses. Considera-se que aproximadamente 35% dos indivíduos da amostra se tornarem inadimplentes no prazo de 70 meses. A técnica utilizada também nos permitiu verificar, considerando um nível de significância de 5%, que há diferença em relação à inadimplência entre homens e mulheres em função do uso de cartões de crédito.

Os resultados demonstraram que o sexo feminino possui maior sobrevivência ao ser comparado com o sexo masculino, entretanto resultado que deve ser considerado as questões de padrão de consumo e comportamento econômico de acordo com os dados da amostra e variáveis analisadas.

Acredita-se que o modelo, eventualmente, salvo as limitações pertinentes a cada tipo de amostra, poderá ser utilizado como base para simulações em que o objetivo é estimar o risco de crédito das carteiras de empréstimos ao consumidor individual, oferecendo suporte gerencial nas decisões de risco.

Em relação às variáveis Limite e Renda Confirmada, utilizadas no estudo, primeiramente analisadas como variáveis da amostra, denominadas de “Cadastro”, aplicadas ao modelo de Regressão Paramétrico demonstraram baixo poder de explicação em relação ao risco de inadimplência em função do uso de cartão de crédito.

Tal resultado demonstrou certa preocupação ao considerar as premissas básicas exigidas no mercado bancário de concessão de crédito. A exigência de comprovação de Renda para obtenção de crédito ao consumo ou financiamentos é obrigatória e em geral, portanto levando-nos ao questionamento da prática atual de concessão de crédito.

Em um segundo momento as mesmas variáveis foram aplicadas ao modelo de regressão semiparamétrico de Cox, modelo de riscos proporcionais com o intuito de analisar uma possível variação na previsão de inadimplência e possível efeito na função de risco do modelo. Novamente os resultados demonstraram pouca influência em relação ao risco, e novamente comprovaram que a variável Limite de Crédito estabelece uma relação inversa ao modelo não podendo, portanto ser considerada como a variável resposta para inadimplência, indicando, todavia menor influencia no aumento de risco de inadimplência.

Por outro lado, o aumento do limite de crédito pode não significar menor risco de inadimplência, do mesmo modo que não explica adimplência, ao considerarmos que o usuário de cartão de crédito poderá aumentar seus gastos na medida em que aumenta o seu limite de crédito levando em consideração o comportamento padrão de consumo.

Por último, confirmando as expectativas as variáveis macroeconômicas Taxa de Juros - Selic, Índice IPCA, e Taxa de Desemprego demonstraram maior poder de explicação em relação ao evento de interesse, neste estudo, inadimplência, comparadas as variáveis da amostra, cadastro. E, portanto, na pratica observada no mercado de crédito as variáveis Cadastro, atualmente utilizadas para comprovação de renda pelos agentes credores pouco oferece proteção frente ao risco de inadimplência incorrido pelo consumidor.

Tal resultado apoia os estudos levantados no referencial teórico, principalmente, nos estudos de Malik e Thomas (2009) onde demonstraram que a intensidade de inadimplência dos consumidores é significativamente influenciada por fatores macroeconômicos.

Em relação à função de risco do modelo, considerado como aumento do risco de *default*, inadimplência, dos usuários de cartão de crédito, os resultados apontaram que a Taxa de Desemprego apresentou o maior poder de explicação em função do risco de inadimplência, sendo a variável mais representativa em relação ao conjunto analisado. Resultado também apresentado nos estudos de Bellotti e Crook (2007), onde demonstram que a taxa de risco aumenta em função da Taxa de Juros e do aumento no nível de desemprego.

A variável IPCA, que verifica as variações dos custos com os gastos das pessoas que ganham de um a quarenta salários mínimos nas regiões metropolitanas do país demonstrou também considerável aumento de risco frente ao evento de inadimplência, demonstrando, portanto aumento em aproximadamente 30% em relação ao risco de inadimplência.

Os resultados apresentados do ponto de vista econômicos demonstram e sugerem que o aumento no índice de inflação irá refletir na perda do poder de pagamento por parte dos consumidores, devido à alta dos preços, e portanto exigindo maior capacidade da variável Limite de Crédito em absorver o impacto deste aumento, caso não haja mudança no padrão de consumo do indivíduo.

Tal fato nos leva a interpretação também, conforme Banco Central (2011) em seu Relatório de Inflação, que mesmo os impactos causados pela taxa de inflação sejam sazonais refletem para o consumidor na elevação de preços dos alimentos *in natura*, de reajustes nas tarifas de transporte público, nos custos com educação, além de influenciar a relação oferta e demanda agregada e a evolução dos preços de serviços, portanto estabelecendo um fator de risco relevante na estimação de risco para carteiras de crédito ao consumo.

Apesar da sua não significância estatística no modelo, como covariável explicativa para inadimplência, resultado a ser investigado conforme sugere Colosimo e Giolo (2006 p.193) alertando para uma possível interpretação de perda de prognóstico na presença de variáveis fortemente associadas no modelo.

E ainda em Tang et al.(2007) confirma-se o entendimento que ao incluir variáveis macroeconômicas em modelos de previsão de risco verifica-se quase que sintomaticamente a leitura subjacente do clima geral econômico do país além de poderem servir de indicadores e sinalizadores de um determinado comportamento econômico.

Embora se observe que apesar dos resultados estatísticos em relação à amostra analisada, o estudo de variáveis que explicam e interferem na expectativa de predição da probabilidade de evento de *default*, e, dada a sua natureza exploratória, tais conclusões se restringem aos dados da amostra que poderá ser considerada limitada do ponto de vista estatístico.

Em relação à disponibilidade de dados, no tratamento estatístico da amostra, salienta-se que devido ao caráter confidencial dos dados em relação às regras brasileiras de sigilo bancário, a obtenção dos dados possa ser um fator de comprometimento, principalmente em relação às variáveis do indivíduo, como por exemplo, Cadastro, sendo um ponto importante em relação aos estudos de crédito ao consumo.

Ainda no que se refere às possíveis interpretações sobre o efeito das variáveis macroeconômicas em modelos de risco de crédito é preciso considerar certas limitações, como por exemplo, características demográficas e mudanças ambientais, pois possuem grande impacto no entendimento sobre variáveis macroeconômicas, entretanto este estudo pode ser considerado um ponto de partida para estudos mais aprofundados neste assunto.

Portanto sugerem-se futuras linhas de investigação em modelos de risco de crédito que possam analisar e ampliar a compreensão sobre a influência das variáveis macroeconômicas em modelos preditivos de risco de crédito ao consumo e conseqüentemente atender as necessidade das instituições financeiras face às necessidades de gestão de risco interno e atendimento aos agentes reguladores. .

Por fim, conclui-se que este estudo cumpriu seu objetivo primário de contribuir para ampliação dos estudos sobre risco de crédito e através dos resultados apresentados ter explicitado as limitações e vantagens em se estimar probabilidade de risco futuro.

6 REFERÊNCIAS

ALLEN L., DELONG G., SAUNDERS A. Issues in the credit risk modeling of retail markets. **Journal of Banking of Finance**, V. 28, n.4, 2004. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0378426603001973>>. Acesso em 04 nov. 2011.

ALTMAN, E. I. Financial ratios, discriminant analysis and the prediction of corporation bankruptcy. **Journal of Finance**, v. 23, n. 4, p. 589-609, 1968.

ALTMAN, E. I., HALDEMAN R.G., NARAYANAN P. Zeta Tin Analysis. A new Model to identify bankruptcy risk of corporation. U.S.A. **Journal of Banking and Finance** 1, pp 29-54. 1977.

ALTMAN, E.I. Predicting Financial Distress of Companies, revisiting the Z-score and Zeta@models, **Journal of Finance**, 2000.

ALTMAN, E.I. Revisiting Credit Scoring Models in a Basel 2 Environment, Ong, M., Credit Rating: Methodologies, Rationale and *Default Risk*, **London Risk Books**, 2002.

ANNIBAL, C.A. Inadimplência do Setor Bancário Brasileiro: uma avaliação de suas medidas. **Banco Central – DEPEC. Trabalhos para Discussão nº 192. 2009**. Disponível em <<http://www.bcb.gov.br/pec/wps/port/wps192.pdf>>. Acesso em 12 set. 2011.

ANDREEVA G. Generic Scoring Models Using Survival Analysis. **The journal of the Operational Research Society**, Vol.57, No.10, pp.1180-1187. 2006.

ANDREEVA G., ANSELL J., CROOK J.N. **Analyzing Revolving Credit. 2003**. Disponível em: <http://www.crc.man.ed.ac.uk/conference/archive/2003/papers/andreeva-ansell-crook.pdf>. Acesso em 20 ago. 2011.

ASSOCIAÇÃO BRASILEIRA DE CARTÕES DE CRÉDITO. **Pagamento eletrônico alavanca o varejo. 2010**. Disponível em:<<http://www.abecs.org.br>>. Acesso em 10 set. 2011.

BASEL COMMITTEE ON BANKING SUPERVISION - BCBS. **International convergence of capital measurement and capital standards**. Basel: BCBS, 2006.

BASEL COMMITTEE ON BANKING SUPERVISION (BCBS), Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework, **BIS**, Basileia, 2004. Disponível em < <http://www.bis.org/publ/bcbs107.htm>>. Acesso em 13 set. 2011.

_____. International convergence of capital measurement and capital standards. **Basel: BCBS**, 1988.

BANCO CENTRAL DO BRASIL. Comunicado 12.746. Procedimentos para a implementação da nova estrutura de capital – **Basiléia II**. Brasília, dezembro 2004.

_____. **Resolução 3.380**. Implementação de Estrutura de Gerenciamento de Risco Operacional, Brasília, junho 2006.

_____. **Circular 2.784**, de 27 de novembro de 1997. Disponível em: < <http://www.bcb.gov.br>.>. Acesso em 11 nov. 2011.

BANCO CENTRAL DO BRASIL - **Relatório sobre a Indústria de Cartões de Pagamentos**, 1ª Ed. 2010.

BANCO CENTRAL DO BRASIL - **Relatório de Inflação**, v.3. 2011.

BANASIK J, CROOK JN, THOMAS LC. **Not if but when will borrowers default**. J Opl Res Soc 50: 1185-1190. 1999.

BAESENS B. GESTEL T.V., STEPANOVA M., POEL D. V.D., VANTHIENEN J.: Neural Network Survival Analysis for Personal Loan Data. **The Journal of the Operational Research Society**, Vol. 56, No. 9, pp.1089-1098. 2005.

BHARATH S. T., SHUMWAY T. Forecasting *Default* with the Merton Distance to *Default* Model .Ross School of Business, University of Michigan Ross School of Business, University of Michigan.**The Review of Financial Studies**. V. 21 n 3, 2008.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Instituto Brasileiro de Geografia e Estatística. Pesquisa Mensal de Emprego-PME. Mar.2012

BELLOTTI T., CROOK J. Modelling and predicting loss given *default* for credit cards. **Quantitative Financial Risk Management Centre**. Working Paper. 2007.

BASTOS J., ROCHA C. Análise de Sobrevivência Conceitos Básicos. Arquivos de Medicina. Vol. 20, Nº 5/6. **Serviço de Higiene e Epidemiologia da Faculdade de Medicina da Universidade do Porto**; Faculdade de Ciências da Universidade de Lisboa, DEIO e CEAUL. **2006**. Disponível em <http://www.scielo.oces.mctes.pt/pdf/am/v20n5-6/v20n5-6a07.pdf>. Acesso em 09 set. 2011.

BRYMAN, A. Social research methods. p. 61-82.2nd. ed. Oxford: **University Press**, 2004.

CAOUCETTE J., ALTMAN E., NARAYANAN P. Managing credit risk: **the next great financial challenge**. New York: John Wiley & Sons, Inc. 1998.

CAPOZZI, Catherine. E How Contributor. **The Economics of Consumer Credit**. Disponível em: <http://www.ehow.com/info_8129443_economics-consumer-credit.html> Acesso em: 13 mar. 2011.

CROUGH M., GALAI D, MARK R. **Risk Management**, McGraw-Hill, New York, 2000.

COLOSIMO, E. A., GIOLO, S. R. **Análise de Sobrevivência Aplicada**. 1 ed. Edgard Blücher Ltda: São Paulo. 2006.

COX, D. R. Regression models and life tables. **Journal of the Royal Statistical Society**, v. 34, n. 2, p. 187-220, 1972.

_____. Partial Likelihood. **Biometrika**, Vol. 62, No. 2, pp. 269-276, 1975.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. 2. ed. Porto Alleger: Artmed, p. 43-63.2007.

EVERITT B., HOTHORN T. **A handbook of statistical analyses using R**. 2nd Ed. CRC Press. Taylor & Francis Group. 2010.

FECOMÉRCIO. Análise da Pesquisa de Endividamento e Inadimplência do Consumidor (**PEIC**). Maio. 2011. São Paulo, SP, 2011. Disponível em <http://www.fecomercio.com.br/?option=com_estudos&view=interna&Itemid=19&id=3903>. Acesso em 22 ago.2011.

FISHER, R. A. The use of multiple measurements in taxonomic problems. **Annals of Eugenics**, 7, p. 179-188. 1936.

FIGINI S., FANTAZZINI D. Random Survival Forests Models for SME Credit Risk Measurement. **Methodology and Computing in Applied Probability**, Forthcoming. 2007. Disponível em :< <http://ssrn.com/abstract=1335856>>. Acesso em 18 jul 2011.

GALINDO J. TAMAYO P. Credit Risk Assessment using Statistical and Machine learning. Basic Methodology and Risk Modeling Applications. **Computational Economics, Kluwer Academic Publishers**, Netherlands, pp.107-13.2000.

GELPI R.M; LABRUYERE F. J. The history of consumer credit: doctrines and practices. **St. Martin's Press**, New York .USA. 2000.

GIL. R. D. Understanding Cox's Regression Model: A Martingale Approach. **Journal of the American Statistical Association**, Vol. 79, No. 386, pp. 441-447. 1984.

GUJARATI, Damodar N. **Basic econometrics**. 3th ed. New York: McGraw-Hill, 2006.

GURTLER M.; HEITHECKER D. Multi-Period *Defaults* and Maturity Effects on economic Capital in a Ratings-Based *Default-Mode* Model – Working Paper series - Braunschweig University of Technology. **Institute for Economics and Business Administration Department of Finance**. 2005.

HAND D.J., Kelly M.G. Lookahead Scorecards for New Fixed Term Credit Products. **The Journal of the Operational Research Society**, Vol.52.No.9, pp.989-996. 2001.

HE.Z.; XIONG W. Rollover Risk and Credit Risk. **Jornal of Finance**. 2010. Disponível em: <<http://www.afajof.org/afa/forthcoming/7848.pdf>> Acesso em 10 nov, 2011.

HIAN, C.K., WEI. C.T., CHWEE. P.G. A two-step method to construct Credit Scoring with data mining Techniques. **International Journal of Business and Information** V.1, n1, pp 96-118. 2006.

JACKSON, P. Basle Committee on banking Supervision. Working Paper no.1. Capital Requirements and Bank Behavior: **The Impact of the Basle Accord**. 1999.

JACKSON, C. H. Multi-state modelling with R: The MSM Package, V6. **Imperial College**: London. 2006.

JACOBS Jr., M.; KARAGOZOGLU, A. K. **Understanding and Predicting Ultimate Loss-Given Default on Bonds and Loans**. Disponível em: <<http://www.fma.org>>. Acesso em: 13 set. 2011.

JACOBS, M. Jr., An empirical study of exposure at *default*, **The Journal Advanced Studies in Finance**. V. 1, N.1, pp. 31-59.2010.

_____. **Validation of economic capital models**: State of the practice, supervisory expectations and results from a bank study, **Journal of Risk Management in Financial Institutions**, V.3, N.4, 334-365.2010.

LI G.H., HAND D.J. Direct versus Indirect Credit Scoring. **The Journal of the Operational Research Society**. Vol. 53, No. 6. pp. 647-654. 2002.

MALIK M., THOMAS L.C. Modelling credit risk of portfolio of consumer loans. **Journal of the Operational Research Society**, 61, p. 411-420. 2009.

MARKOWITZ, H.M. Portfolio Selection. **The Journal of Finance** 7, 1 pp. 77–91. 1952.

MESTER, L.J. What's the point of credit scoring? **Business Review** .September. 3-16. 1997.

MERTON, R. C. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. **Journal of Finance** 29:449–70. 1974.

PESARAN M.; HASHEM. M. SCHUERMAN T. TREUTLER B. J. WEINER. S.M. Macroeconomic Dynamics and Credit Risk: A Global Perspective. **Journal of Money, Credit, and Banking**, V. 38, N 5, pp. 1211-1261, 2006.

PINDYCK, Robert S, RUBINFELD, Daniel L. **Econometria**: Modelos & Previsões. Rio de Janeiro: Elsevier, 2004. 4.ed

POEL V. D., LARIVIERE B. Customer Attrition analysis for financial services using Proportional Hazard models, working paper, Department of Marketing, **Ghent University**, Ghent, 2003.

KLEINBAUM D.G.; KLEIN M. Survival Analysis. A Self-Learning Text Springer Science **Business Media**, Inc. 2nd. ed. 1996

KACHMAN, S. Applications in survival analysis. **Journal Animal Science**, n.77, p.147-153.1999.

ROSCH D.; SCHEULE H. Forecasting Retail Portfolio Credit Risk. **Journal of Risk Finance**, 5 (2), 16-32. 2004.

RIPLEY, B.D. Neural networks and methods for classification. **J.R. Statist. Soc. B**, 56, 409-456.1994.

RELATÓRIO DE ESTABILIDADE FINANCEIRA, **Banco Central do Brasil**, volume 10, nº 1. 2011. Disponível em< http://www.bcb.gov.br/htms/estabilidade/2011_04/refP.pdf>. Acesso em 12 ago 2011.

SAUNDERS, Anthony. Credit Risk Measurement: New Approaches to Value-at-Risk and other Paradigms. **Wiley John & Sons**, 1. Ed. 1999.

SCHUERMANN, Til. What Do We Know About Loss Given Default? The Wharton Financial Institution Center, 2004.

STEPANOVA. M., THOMAS. L.C. Survival analysis for personal loan data. **Opl Res** 50: 277-28. 2002.

TANG L., THOMAS L.C. THOMAS S., BOZZETTO J.F. School of Management, University of Southampton, Southampton, UK. It's the economy stupid: modeling financial product purchases. **International Journal of Bank Marketing** Vol. 25 No. 1, pp. 22-38. 2007.

THOMAS L.C., OLIVER R. W., HAND D. J. A Survey of the Issues in Consumer Credit Modelling Research. **The Journal of the Operational Research Society**, Vol. 56, No. 9 pp.1006. 2005.

THOMAS, L. C.; David B.E., Credit Scoring and Its Applications. Cambridge, **Synopses & Reviews**. 2002.

THOMAS, L. C. Consumer credit models: pricing, profit, and portfolios. **Oxford University Press**. 2009.

TSIATIS, A.A. A Large Sample Study of Cox's Regression Model. **The Annals of Statistics**, Vol. 9, No. 1, pp 93-108, 1981.

VASICEK O. A. The Distribution of Loan Portfolio Value. **Published in Risk**, December. 2002.

VERSTRAETEN G., POEL D.V.D. The Impact of Sample Bias on Consumer Credit Scoring Performance and Profitability. **The Journal of the Operational Research Society**, Vol 56, No.8.pp. 981-992. 2005.

WHALEN G.A; A Proportional Hazards Model of bank failure: An examination of its usefulness as Early Warning tool, **Economic Review**, issue Q I, pages 21-31, 1991.

WILSON, Thomas C. Credit Portfolio Risk. FRBNY ECONOMIC POLICY REVIEW / OCTOBER 1998.

Softwares Estadísticos:

R Development Core Team (2008). R: **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

STATA CORPORATION. **Stata Statistical Software version 8.0**. [S.I.]: College Station, TX/ Stata Corporation, 2003.

S586e Silva, Sandra Almeida

Estudo de risco de crédito em operações de cartão de crédito usando variáveis macroeconômicas e técnicas de análise de sobrevivência / Sandra Almeida Silva - 2012.

86f. : il., 30 cm

Dissertação (Mestrado em Administração de Empresas) – Universidade Presbiteriana Mackenzie, São Paulo, 2012.

Orientação: Prof. Dr. Herbert Kimura

Bibliografia: f. 81-86

1. Análise de sobrevivência. 2. Inadimplência. 3. Variáveis macroeconômicas. I. Título.

CDD 658.88