

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

Rodrigo Heldt Silveira

O VALOR FUTURO DE CADA CLIENTE:
ESTIMAÇÃO DO *CUSTOMER LIFETIME VALUE*

Porto Alegre

2014

Rodrigo Heldt Silveira

O VALOR FUTURO DE CADA CLIENTE:
ESTIMAÇÃO DO *CUSTOMER LIFETIME VALUE*

Dissertação apresentada no Programa de Pós-Graduação em Administração da Universidade Federal do Rio grande do Sul (PPGA/UFRGS) como requisito parcial para obtenção de grau de Mestre em Administração – Ênfase em *Marketing*.

Professor Orientador: Dr. Fernando Bins Luce

Porto Alegre

2014

CIP - Catalogação na Publicação

Silveira, Rodrigo Heldt

O valor futuro de cada cliente: estimação do customer lifetime value / Rodrigo Heldt Silveira. -- 2014.

135 f.

Orientador: Fernando Bins Luce.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Escola de Administração, Programa de Pós-Graduação em Administração, Porto Alegre, BR-RS, 2014.

1. Customer Lifetime Value. 2. Customer Equity. 3. Segmentação de Clientes. 4. Marketing Metrics. 5. Estimação Bayesiana. I. Luce, Fernando Bins , orient. II. Título.

Rodrigo Heldt Silveira

O VALOR FUTURO DE CADA CLIENTE:
ESTIMAÇÃO DO *CUSTOMER LIFETIME VALUE*

Dissertação apresentada no Programa de Pós-Graduação em Administração da Universidade Federal do Rio grande do Sul (PPGA/UFRGS) como requisito parcial para obtenção de grau de Mestre em Administração – Ênfase em *Marketing*.

Professor Orientador: Dr. Fernando Bins Luce

Aprovada em 21 de agosto de 2014

BANCA EXAMINADORA:

Prof. Dr. José Afonso Mazzon

Prof. Dr. Luiz Antônio Slongo

Prof. Dr. João Luiz Becker

Porto Alegre
2014

AGRADECIMENTOS

A confiança na relevância teórica e prática do estudo realizado foi o que fez com que os desafios impostos, principalmente pelo método adotado, fossem superados. O processo de desenvolvimento do presente estudo exigiu muito estudo e perseverança para superar as dúvidas, possibilitando as descobertas e aprendizados necessários. Esse caminho, entretanto, só pôde ser trilhado devido ao apoio e auxílio de um conjunto muito especial de pessoas que, de diferentes formas, foram essenciais para o atingimento dos objetivos propostos.

Agradeço, em primeiro lugar, ao meu orientador, Prof. Fernando Bins Luce, presente de forma marcante não só durante o presente estudo, mas durante toda a minha trajetória no curso de mestrado. Agradeço também a participação importante no trabalho realizado do Prof. José Afonso Mazzon e do Prof. João Luiz Becker. Sou muito grato também aos ensinamentos durante o curso do Prof. Luiz Antônio Slongo, da Prof. Cristiane Pizzutti dos Santos, do Prof. Carlos Alberto Vargas Rossi e do Prof. Walter Meucci Nique, tendo esse último dado apoio e incentivo importantes na etapa anterior a minha entrada no curso de mestrado.

Agradeço, em especial, à Doutoranda em Marketing Cleo Schmitt Silveira, uma verdadeira parceira durante a realização do presente estudo, com quem pude aprender muito, bem como pude contar com seu apoio desde o primeiro dia de trabalho. Agradeço, também em especial, ao Mestre em estatística Filipe Jaeger Zabala, que foi outro verdadeiro companheiro durante a realização do trabalho e com quem pude ter discussões intelectualmente valiosas. Sou também especialmente grato ao Doutorando em Tecnologia da Informação Camilo Bornia, quem proporcionou auxílio essencial sobre gerenciamento da base de dados utilizada e sobre programação na linguagem *SQL*, e ao meu colega de mestrado Rodrigo Silva, quem também auxiliou na tarefa de gerenciamento da base de dados utilizada.

Sou muito grato a todos os meus colegas de mestrado. Durante todo o curso fomos sempre companheiros, tanto naquelas fases mais críticas quanto nos momentos de confraternização. Essa agradável convivência gerou laços de amizade que certamente serão cultivados.

Por fim, agradeço a minha namorada, quem sempre me apoiou e, até nos momentos em que não podia compreender os problemas técnicos enfrentados, me ouvia e me dava conselhos; e ao meu pai, à minha mãe e aos meus irmãos, que são aqueles que mais me apoiaram não só nessa etapa, mas em cada passo da minha vida, que me ensinaram os alicerces que hoje fazem com que eu consiga trilhar meu caminho. A todos esses devo o meu amor, franco e verdadeiro.

**“A verdadeira viagem de descobrimento
não consiste em procurar novas paisagens,
mas em ter novos olhos.”**

Marcel Proust

**“A mente que se abre a uma nova ideia
jamais voltará ao seu tamanho original.”**

Albert Einstein

RESUMO

A capacidade de o *marketing* mensurar e comunicar o valor de suas atividades e investimentos tem sido uma das prioridades de pesquisa na área nos últimos anos. Para atingir esse objetivo, a capacidade de mensurar adequadamente os ativos de *marketing*, como o *Customer Lifetime Value* e, de forma agregada, o *Customer Equity*, torna-se essencial, pois esses ativos são considerados os elementos capazes de traduzir em valores monetários o resultado dos diversos investimentos realizados pela área de *marketing*. Diante da mensuração desses valores, é possível o planejamento e a realização de ações mais precisas por parte dos profissionais de *marketing*. Sendo assim, no presente estudo objetivou-se construir e aplicar um modelo de estimação de *Customer Lifetime Value* no modo *bottom-up* (individual por cliente) em uma amostra de clientes de uma empresa do setor de serviços financeiros. O modelo bayesiano hierárquico aplicado, com três regressões estruturadas conforme o modelo *Seemingly Unrelated Regressions (SUR)* (ZELNER, 1971), foi construído a partir dos trabalhos de Kumar *et al.* (2008), Kumar e Shah (2009) e Cowles, Carlin e Connet (1996). Os resultados evidenciaram (1) que o modelo foi capaz de estimar com consistência o valor futuro de 84% dos clientes analisados; (2) que esse valor estimado traduz o potencial de rentabilidade que pode ser esperado futuramente para cada cliente; (3) que a base de clientes pode ser segmentada a partir do *Customer Lifetime Value*. Diante do conhecimento do valor futuro de cada cliente, se vislumbrou possibilidades de ações que tragam melhorias para gestão de clientes tradicionalmente utilizada, principalmente no que diz respeito à alocação dos recursos de *marketing*.

Palavras-chave: *customer lifetime value*, *customer equity*, segmentação de clientes, estimação bayesiana, imputação de *missing values*.

ABSTRACT

The marketing capacity to measure and to communicate the value resultant of its activities and investments has been one of the area top research priorities in the last few years. In order to achieve this objective, the capacity to appropriately measure the marketing assets, as the Customer Lifetime Value and, in aggregate form, the Customer Equity, has been pointed out as essential, because this assets are considered elements capable of translating the result of marketing investments into monetary values. Given the measurement of those values, marketers become able to plan and take more precise actions. Thus, the objective of present study is to build and test a bottom-up Customer Lifetime Value estimation model to a sample of customers from a company of finance services. The bayesian hierarchical model, composed of three regressions structured according to the Seemingly Unrelated Regressions (SUR) model (ZELNER, 1971), was built from the works of Kumar *et al.* (2008), Kumar and Shah (2009) and Cowles, Carlin and Connet (1996). The results show that (1) the model was capable to estimate with consistency the future value of 84% of the analyzed customers; (2) this estimated future values indicate the potential profitability of each customer; (3) the customer base can be segmented from the Customer Lifetime Value. Given the knowledge obtained about the future value of each customer and the segments established, several actions that can bring improvements to the traditional way of managing customers were suggested, in special those concerning marketing resource allocation.

Keywords: customer lifetime value, customer equity, customer segmentation, bayesian estimation, data augmentation.

LISTA DE FIGURAS

Figura 1 – Cadeia da Produtividade de <i>Marketing</i> (RUST <i>et al.</i> , 2004)	22
Figura 2 – Cadeia da Produtividade de <i>Marketing</i> (LEHMANN, 2004)	24
Figura 3 – Modo <i>top-down</i> para mensurar o <i>CLV</i>	29
Figura 4 – Modo <i>bottom-up</i> para mensurar o <i>CLV</i>	30
Figura 5 – Gráficos de densidade e traço da cadeia para a variável custos de <i>marketing lag 1</i> – Regressão 1	91
Figura 6 – Gráficos de densidade e traço da cadeia para a variável custos de <i>marketing lag 2</i> – Regressão 1	91
Figura 7 – Gráficos de densidade e traço da cadeia para a variável MC total <i>lag 1</i> – Regressão 1	92
Figura 8 – Gráficos de densidade e traço da cadeia para a variável <i>net</i> – Regressão 1 ...	92
Figura 9 – Gráficos de densidade e traço da cadeia para a variável <i>cross-buying</i> – Regressão 2	92
Figura 10 – Gráficos de densidade e traço da cadeia para a variável <i>buy lag1</i> – Regressão 2	92
Figura 11 – Gráficos de densidade e traço da cadeia para a variável <i>buy lag2</i> – Regressão 2	93
Figura 12 – Gráficos de densidade e traço da cadeia para a variável frequência de compra – Regressão 2	93
Figura 13 – Gráficos de densidade e traço da cadeia para a variável <i>cross-buying</i> – Regressão 3	93
Figura 14 – Gráficos de densidade e traço da cadeia para a variável MC total <i>lag1</i> – Regressão 3	93
Figura 15 – Gráficos de densidade e traço da cadeia para a variável MC total <i>lag2</i> – Regressão 3	94
Figura 16 – Gráficos de densidade e traço da cadeia para a variável total de dias com transação – Regressão 3	94
Figura 17 – Gráficos de densidade e traço da cadeia para a variável <i>net</i> – Regressão 3.	94
Figura 18 – Segmentação da base de clientes a partir o <i>CLV</i> - modelo aplicado.....	97
Figura 19 – Segmentação da base de clientes a partir o <i>CLV</i> - modelo ingênuo	99

LISTA DE TABELAS

Tabela 1 – Pressupostos acerca da interface entre <i>marketing</i> e finanças.....	20
Tabela 2 – Cálculo da margem de contribuição por categoria de produto.....	60
Tabela 3 – Variáveis consideradas.....	64
Tabela 4 – Análises Univariadas – variáveis de razão.....	83
Tabela 5 – Análises Univariadas – variáveis categóricas.....	85
Tabela 6 – Análises de correlação.....	87
Tabela 7 – Parâmetros estimados.....	89
Tabela 8 – Resultados para os testes de convergência.....	90

SUMÁRIO

1 INTRODUÇÃO	13
1.1 DELIMITAÇÃO DO TEMA, DEFINIÇÃO DO PROBLEMA E JUSTIFICATIVA 14	
1.2 OBJETIVOS DA PESQUISA	16
1.2.1 Objetivo Geral	17
1.2.2 Objetivos Específicos	17
1.3 ORGANIZAÇÃO DO TRABALHO.....	17
2 REFERENCIAL TEÓRICO	17
2.1 <i>MARKETING</i> RESPONSÁVEL PELA MENSURAÇÃO DA PRÓPRIA PRODUTIVIDADE	18
2.2 <i>CUSTOMER LIFETIME VALUE (CLV)</i> e <i>CUSTOMER EQUITY (CE)</i>	25
2.3 MODELOS DE ESTIMAÇÃO DO <i>CUSTOMER LIFETIME VALUE (CUSTOMER EQUITY)</i>	33
2.3.1 Jain e Singh (2002)	34
2.3.2 Berger e Nasr (1998)	35
2.3.3 Blattberg e Deighton (1996)	36
2.3.4 Rust, Lemon e Zeithaml (2004)	37
2.3.5 Gupta, Lehmann e Stuart (2004)	39
2.3.6 Schmittlein, Morrison e Colombo (1987) / Schmittlein e Peterson (1994)	40
2.3.7 Fader, Hardie e Lee (2005)	41
2.3.8 Modelos <i>RFM</i> (Em Inglês: <i>Recency, Frequency and Monetary</i>)	42
2.3.9 Borle, Singh e Jain (2008)	43
2.3.10 Venkatesan, Kumar e Bohling (2007)	44
2.3.11 Kumar <i>et al.</i> (2008)	46
2.3.12 Kumar e Shah (2009)	47
3 MÉTODO	56
3.1 DADOS UTILIZADOS.....	56
3.1.1 Descrição da empresa	56
3.1.2 Base de dados	57
3.2 DESCRIÇÃO DO MODELO CONSTRUÍDO E APLICADO	63
3.3 ESTIMAÇÃO DOS PARÂMETROS	71
3.4 ANÁLISES DE DIAGNÓSTICO DE CONVERGÊNCIA	76
3.4.1 Critério de Geweke	76

3.4.2 Critério de Heidelberger e Welch	77
3.4.3 Critério <i>Half-Width</i>	78
3.5 CÁLCULO DO <i>CUSTOMER LIFETIME VALUE</i> E DO <i>CUSTOMER EQUITY</i>	78
3.6 SEGMENTAÇÃO DA BASE DE CLIENTES A PARTIR DO <i>CUSTOMER LIFETIME VALUE</i>	81
3.7 VERIFICAÇÃO DA CAPACIDADE DE PREVISÃO DO MODELO APLICADO...	82
4 RESULTADOS	82
4.1 ANÁLISES DESCRITIVAS	83
4.1.1 Análises Univariadas.....	83
4.1.2 Análises de Correlação.....	85
4.2 ESTIMAÇÃO DOS PARÂMETROS	86
4.3.1 Análises de diagnóstico de convergência	90
4.3 CÁLCULO DO <i>CUSTOMER LIFETIME VALUE</i> E DO <i>CUSTOMER EQUITY</i>	95
4.4 SEGMENTAÇÃO DA AMOSTRA DE CLIENTES A PARTIR DO <i>CLV</i>	96
4.5 VERIFICAÇÃO DA CAPACIDADE DE PREVISÃO DO MODELO APLICADO...	99
4.5.1 Comparação com o modelo ingênuo.....	99
4.5.2 Análise cliente a cliente	101
4.5.3 Análise da previsão do período subsequente	103
5 CONCLUSÃO.....	104
5.1 CONSIDERAÇÕES FINAIS	104
5.2 LIMITAÇÕES DA PESQUISA E SUGESTÕES PARA PESQUISAS FUTURAS...	106
REFERÊNCIAS	107
ANEXOS	117
ANEXO A – MÉDIAS E DESVIOS PADRÃO DOS INTERCEPTOS	117
ANEXO B – RESULTADOS DOS CRITÉRIOS DE CONVERGÊNCIAS PARA OS INTERCEPTOS	121
ANEXO C - <i>CLV</i> ESTIMADO PARA CADA CLIENTE	125
ANEXO D - ANÁLISE CLIENTE A CLIENTE PARA VERIFICAÇÃO DA CAPACIDADE DE PREVISÃO	128

1 INTRODUÇÃO

Entre a década de 50 e a década de 70, diante do baixo nível de desenvolvimento dos mercados, o *marketing* foi capaz de contribuir de forma determinante para o crescimento das empresas e dos mercados como um todo (KUMAR, 2004). A partir da década de 80, porém, os mercados passaram a ficar mais complexos e novas formas de organização surgiram, incluindo alianças estratégicas e redes (WEBSTER, 1992). Com isso, surgiu a necessidade de novas contribuições da área de *marketing* (WEBSTER 1988, 1992, 1994, 2005).

Diante disso, nos anos 90, os 4Ps (produto, preço, promoção e praça), que outrora foram essenciais para o desenvolvimento das organizações, perderam definitivamente a sua força como componente estratégico (DAY e MONTGOMERY, 1999; KUMAR, 2004). Assim, o *marketing* como motor do desenvolvimento acabou ficando do seu potencial (GÖK e HACIOGLU, 2010; KUMAR, 2004). Ironicamente, ao mesmo tempo em que entrou em declínio, essa função nunca foi tão necessária para as organizações (KUMAR, 2004). Desde então, os esforços na área de *marketing* têm objetivado encontrar formas de retomar seu papel estratégico junto à mesa da diretoria.

Entre as causas dessa perda de espaço da área de *marketing* dentro das organizações, a falta de capacidade de mensurar a produtividade de suas ações e de relacionar esses resultados com a geração de valor para o acionista é vista como um dos principais desafios que esta área deve enfrentar para voltar a ter participação estratégica (RUST *et al.*, 2004; WEBSTER, MALTER e GANESAN, 2005). Para Sheth e Sisódia (1995a; 1995b), as ações de *marketing*, na época, não eram adequadamente mensuradas e, assim, sem métricas confiáveis, melhorias nos níveis de eficiência da área seriam extremamente difíceis de serem atingidos. Diante disso, em 1999, com o chamado feito no *MSI Workshop/Conference on Marketing Metrics* (1999), o desenvolvimento de formas de mensuração de métricas de *marketing* definitivamente passou a fazer parte da agenda de pesquisa em *marketing*. Passados 14 anos, esforços nesse sentido foram realizados, no entanto avanços ainda são necessários, o que é comprovado pela decisão do *Marketing Science Institute (MSI)* ao citar a capacidade de mensuração dos seus resultados de *marketing* e o impacto de suas ações na *performance* global da empresa como uma das prioridades de pesquisa durante o biênio 2008 – 2010 (MSI, 2011). Para o biênio 2014-2016, novamente a capacidade de o *marketing* mensurar e comunicar o valor de suas atividades e investimentos foi elencada entre as prioridades de pesquisa pelo *Marketing Science Institute (MSI)* (MSI, 2014).

Para atingir esse objetivo, a capacidade de mensurar adequadamente os ativos de *marketing* (clientes, marcas, canais de distribuição, entre outros) torna-se essencial, pois esses ativos são considerados os elementos capazes de traduzir em valores monetários o resultado dos diversos investimentos realizados pela área de *marketing* (LEHMANN, 2004; RUST *et al.*, 2004). Diante da mensuração desses valores, é possível o planejamento e a realização de ações mais precisas por parte dos profissionais de *marketing*. Além disso, torna-se possível vislumbrar, controlando-se a influência de outras variáveis, uma comparação entre variações no valor desses ativos com variações no valor da firma ou outras variáveis que representem a riqueza do acionista.

A mensuração de um desses ativos de *marketing* – os clientes – por meio da estimação do valor de um cliente específico da base de clientes de uma organização, representado pelo *Customer Lifetime Value (CLV a partir daqui)*, e da estimação do valor total de todos os clientes atuais e futuros de uma organização, representado pelo *Customer Equity (CE a partir daqui)*, tem-se mostrado uma forma eficiente e eficaz de cumprir essa tarefa. Para Kumar e Shah (2009), a ideia de gerenciar os clientes com base no *CLV* (ou *CE*, de forma agregada) surgiu como uma das formas mais eficientes de fazer negócios. Gupta e Lehmann (2006), por sua vez, entendem que o *CLV* está remodelando a forma como se gerencia para maximizar lucros. Os modelos de *CLV* são, assim, poderosas ferramentas para promover a maximização dos retornos dos investimentos de *marketing* e para guiar a alocação adequada do orçamento de *marketing* (VILLANUEVA e HANSSENS, 2007). Existe, ainda, grande otimismo acerca da possibilidade de relacionar o *CE* com o valor gerado para o acionista (GUPTA e LEHMANN, 2006). Diante disso, diversos foram os esforços a partir da década de 90 para construir modelos mais eficientes de *CLV*. As propostas partem de modelos determinísticos bastante singelos e chegam a modelos estocásticos sofisticados tanto em relação à escolha das variáveis consideradas para o cálculo do *CLV* quanto em relação ao método de estimação utilizado.

Para melhor compreender os anseios deste trabalho, essa introdução apresenta, em sua sequência, a delimitação do tema, definição do problema, justificativa e os objetivos desta dissertação.

1.1 DELIMITAÇÃO DO TEMA, DEFINIÇÃO DO PROBLEMA E JUSTIFICATIVA

Diante de um ilimitado orçamento de *marketing*, os gerentes podem decidir contatar todos os seus clientes a cada período. Essa estratégia minimizaria o erro de tipo I, que é não contatar um cliente que poderia ter uma receita potencial a oferecer. No entanto minimizar tal erro também maximiza o erro de tipo II. Um erro de tipo II é contatar um cliente que não está pronto para comprar e é caro por afetar de forma adversa tanto as receitas (*top-line*) quanto os lucros (*bottom-line*). Quando frente a um orçamento limitado de *marketing*, o *trade-off* entre o erro de tipo I e o erro de tipo II torna-se ainda mais relevante e os gerentes são forçados a priorizar as suas estratégias de comunicação na direção daqueles clientes que, conforme esperado, deverão trazer o maior crescimento aos fluxos de caixa (VENKATESAN, KUMAR e BOHLING, 2007).

A partir disso, o *CLV* é visto como uma métrica adequada para promover a aquisição e retenção dos clientes ‘certos’ em uma gestão de relacionamento com clientes (VENKATESAN e KUMAR, 2004), o que indica que o conceito de *CLV* está baseado na visão de que a empresa deve buscar estabelecer relacionamentos com clientes, no entanto não com todo e qualquer cliente, mas com aqueles que geram maior valor. Por isso, muitas empresas estão atualmente priorizando a identificação dos seus clientes mais rentáveis e buscando intensificar a relação com eles. O centro desse movimento é o *CLV*, que pode ser usado para guiar as atividades de aquisição e retenção, além de poder ser usado de forma agregada como uma métrica de valor da firma ou de um segmento de clientes (BLATTBERG, MALTHOUSE e NESLIN, 2009).

Diante da importância flagrante do *CLV*, tanto para a pesquisa em *marketing* quanto para a recuperação do espaço perdido pelo *marketing* dentro das empresas, muitos acadêmicos da área têm despendido esforços para mensurar de forma adequada essa métrica. Os conceitos de *CLV* e também de *CE* são relativamente simples, no entanto a complexidade do tema está relacionada com o método utilizado para estimar esses valores e as variáveis que os determinam. Sendo assim, a maior parte da literatura sobre o tema diz respeito aos esforços para construir modelos melhores para estimar os valores futuros dessas variáveis (GUPTA e LEHMANN, 2006). Alguns dos modelos existentes são do tipo *bottom-up*, ou seja, a partir de uma base de dados que reúna informações precisas sobre cada cliente de uma dada empresa, busca-se estimar o valor vitalício de cada cliente, levando em consideração a heterogeneidade existente entre esses clientes. Modelos assim proporcionam que todo o potencial para a gestão de clientes que emana do *CLV* seja posto em prática.

Entre esses, os modelos de Kumar *et al.* (2008) e de Kumar e Shah (2009), que são intimamente relacionados, uma vez que o segundo traz incrementos em relação ao primeiro, destacam-se por sua capacidade de estimar o valor vitalício dos clientes em nível individual (*bottom-up*) por meio de estatística bayesiana, a qual é ideal para a aplicação em problemas nos quais a decisão precisa ser feita com a modelagem de incerteza. Nesses modelos, os autores utilizam variáveis demográficas e transacionais relacionadas a cada cliente, que são utilizadas para a estimação das variáveis dependentes: probabilidade de compra, margem de contribuição e número de contatos de *marketing* que a empresa realiza com cada cliente. Essas variáveis dependentes são, então, utilizadas para estimar os fluxos de caixa futuros gerados por cada cliente, possibilitando uma sofisticada previsão do *CLV* em nível individual.

Sendo assim, devido à contribuição desses modelos para a segmentação de clientes, para a alocação eficiente e eficaz dos recursos de *marketing* e pelo avanço proporcionado à tarefa de ligar o desempenho de *marketing* ao valor para o acionista, o objetivo principal do presente estudo é construir e aplicar um modelo de estimação de *CLV* a partir dos modelos de Kumar *et al.* (2008) e de Kumar e Shah (2009) para uma amostra de clientes de uma empresa de serviços financeiros em atuação no mercado nacional. ,

A partir do cumprimento da tarefa de aplicar o modelo construído, constatou-se que, quando estimado de forma consistente, o *CLV*, conforme a literatura sustenta, traduz a expectativa futura de rentabilidade de cada cliente, que passa a ser facilmente identificada, proporcionando uma nova gama de possibilidades para os gestores de *marketing* conduzirem o gerenciamento de seus clientes. Uma dessas possibilidades é a definição de segmentos de clientes a partir de seus *CLVs*, o que torna mais objetiva a tarefa de planejar a alocação dos esforços de vendas, a forma de atendimento a cada cliente, as regras de descontos concedidos, a remuneração variável da força de vendas, entre diversas outras possibilidades.

Sendo assim, o problema de pesquisa é definido da seguinte maneira: Quais os resultados da aplicação de um modelo de estimação do *Customer Lifetime Value* individual por cliente em uma amostra de clientes de uma empresa do setor de serviços financeiros?

1.2 OBJETIVOS DA PESQUISA

O objetivo geral e os objetivos específicos desta dissertação são apresentados nos tópicos a seguir.

1.2.1 Objetivo Geral

Aplicar modelo de estimação de *Customer Lifetime Value* individual por cliente em uma amostra de clientes de uma empresa do setor de serviços financeiros.

1.2.2 Objetivos Específicos

Este estudo apresenta os seguintes objetivos específicos:

- Obter e organizar as informações da base de dados da empresa;
- Operacionalizar o modelo:
 - Determinar os componentes do modelo e suas interrelações;
 - Construir o código em linguagem R para o cálculo do modelo;
 - Realizar as análises de diagnóstico sobre os parâmetros calculados;
 - Realizar a estimação das variáveis que compõem o cálculo do *CLV*;
- Calcular o *Customer Lifetime Value* e o *Customer Equity*, sendo esse último baseado na suposição de que a base de clientes é formada apenas pela amostra analisada;
- Segmentar a base de clientes a partir do *Customer Lifetime Value*.

1.3 ORGANIZAÇÃO DO TRABALHO

A apresentação da dissertação está estruturada em mais quatro capítulos (2-5). No primeiro apresenta-se o referencial teórico, que trata da necessidade de o *marketing* ser responsável pela mensuração de sua produtividade, bem como da literatura sobre *Customer Lifetime Value* e *Customer Equity*. No capítulo 3, o método que foi utilizado para aplicar o modelo utilizado é apresentado. No capítulo 4, os resultados do modelo aplicado, bem como as estimações do *CLV* e do *CE* para a amostra analisada e a segmentação dos clientes com base no *CLV* são apresentados. Por fim, no capítulo 5, as limitações da pesquisa e as sugestões para pesquisas futuras são apresentadas, bem como as considerações finais são tecidas abordando-se as conclusões do estudo conduzido.

2 REFERENCIAL TEÓRICO

No presente capítulo será apresentado o referencial teórico que sustenta a necessidade de o *marketing* ser responsável pela mensuração de valor criado por suas atividades e investimentos. A seguir, é apresentado o referencial teórico sobre *CLV* e *CE*, métricas que são alternativas importantes para a mensuração da produtividade de *marketing*.

2.1 *MARKETING* RESPONSÁVEL PELA MENSURAÇÃO DA PRÓPRIA PRODUTIVIDADE

Embora algumas pesquisas mostrem argumentos contrários (HOMBURG, WORKMAN e KROHMER, 1999), muitos autores sustentam que o *marketing* teve sua participação estratégica dentro das empresas reduzida, de forma que alguns autores afirmam que ele perdeu seu lugar na mesa de reuniões da diretoria (DAY, 1992; KUMAR, 2004; WEBSTER, 1992, 1994, 2002; WEBSTER, MALTER e GANESAN, 2003). Ambler (2003) cita que pesquisas realizadas demonstram que, em grandes empresas, apenas 10% do tempo gasto em reuniões de seus executivos é dedicada ao *marketing*. Entre os caminhos vislumbrados para que essa situação seja modificada, entende-se que o *marketing* precisa passar a ser responsável pela mensuração da própria produtividade, buscando encontrar métodos para comprovar, mensurar o impacto de suas ações no desempenho global da empresa e, assim, recuperar seu espaço na mesa de reuniões da diretoria (DOYLE, 2000; MORGAN, CLARK e GOONER, 2002; RUST, LEMON e ZEITHAML, 2004; SHETH e SISODIA, 1995a; 1995b).

A mensuração da produtividade de *marketing* é o problema mais urgente que a gestão de *marketing* enfrenta ao objetivar recuperar seu lugar na mesa da diretoria (WEBSTER, MALTER e GANESAN, 2005). Sheth e Sisódia (1995a; 1995b), ao publicarem dois artigos em sequência em 1995, trouxeram à tona a necessidade de a área de *marketing* ser *accountable*, ou seja, responsável por mensurar os resultados de suas atividades. Para esses autores, as ações de *marketing*, na época, não eram adequadamente mensuradas e, sem métricas confiáveis, melhorias nos níveis de eficiência da área seriam extremamente difíceis de serem atingidas. No entanto, foi em 1999, com o *MSI Workshop/Conference on Marketing Metrics* (1999), que o desenvolvimento de formas de mensuração de métricas de *marketing* definitivamente passou a fazer parte da agenda de pesquisa em *marketing*. Nessa conferência, os acadêmicos Eugene Anderson, Donald R. Lehmann, Sunil Gupta, Rajendra K. Srivastava e Jerry Wind apresentaram trabalhos e perspectivas acerca do Valor do cliente, Valor da Marca,

Métricas do *Mix de Marketing*, Novas Métricas Financeiras e Valoração do *E-Commerce*, respectivamente. Passados 14 anos, esforços nesse sentido foram realizados, no entanto o *marketing* ainda necessita muito de avanços nessa área. A capacidade de o *marketing* mensurar os seus resultados e o impacto de suas ações na *performance* global da empresa foi citada pelo *Marketing Science Institute* (MSI) como uma das prioridades de pesquisa em *marketing* durante o biênio 2008 – 2010 (MSI, 2011). Para o biênio 2014-2016, novamente a capacidade de o *marketing* mensurar e comunicar o valor de suas atividades e investimentos foi elencada entre as prioridades de pesquisa pelo *Marketing Science Institute* (MSI) (MSI, 2014).

Frequentemente o *marketing* tende a se focar em crescimento de vendas e *market-share*, falhando em reconhecer o impacto das decisões de *marketing* em variáveis como níveis de estoque, necessidade de capital de giro, custos de financiamentos, razão entre dívidas e patrimônio e preço das ações. Assumir que tais fatores são puramente uma responsabilidade de finanças é ser culpado por uma forma de miopia de *marketing* não menos danosa do que aquela originalmente identificada por Levitt (1960) (ANDERSON, 1979).

Com base nessa realidade, Srivastava, Shervani e Fahey (1998) buscaram lançar novos pressupostos que promovam uma interface entre *marketing* e finanças, o que faria com que o objetivo principal do *marketing*, embora devesse continuar buscando o tradicional objetivo de sucesso de um produto em seu mercado, fosse contribuir para o incremento dos retornos dos acionistas. Diante disso, esses autores estabeleceram um conjunto de mudanças que os profissionais de *marketing* deveriam incorporar em suas ações para promover a pretendida interface entre *marketing* e finanças (ver Tabela 1).

Essa perspectiva, também aceita por diversos outros autores, como Day e Fahey (1988), pressupõe, então, que o novo papel do *marketing* é criar e gerir os ativos baseados em mercado, tais como clientes (*Customer Equity*), marcas (*Brand Equity*) e canais de distribuição, para entregar valor aos acionistas. Além disso, outras mudanças em relação à maneira tradicional do *marketing* atuar também devem ocorrer na mentalidade dos profissionais de *marketing*. Entre elas, cabe destacar a busca da expansão do conjunto de métricas de desempenho das atividades de *marketing* (SRIVASTAVA, SHERVANI e FAHEY, 1998). Para esses autores, os profissionais de *marketing* deveriam ir além das tradicionais métricas de *marketing* (exemplo: volume de vendas, *market-share* e margem

bruta) e também considerar métricas financeiras, como valor presente líquido dos fluxos de caixa futuros e o valor para o acionista.

Tabela 1: Pressupostos acerca da interface entre *marketing* e finanças

	Pressupostos Tradicionais	Pressupostos Emergentes
Objetivo do <i>marketing</i>	Criar valor para os consumidores. Vencer no mercado de produtos.	Criar e gerenciar ativos baseados em mercado para entregar valor ao acionista.
Relação entre <i>marketing</i> e finanças	Resultados positivos de produto-mercado traduzidos em resultados financeiros positivos.	Interface entre <i>marketing</i> e finanças deve ser gerenciada sistematicamente.
Perspectiva sobre os clientes e os canais	O objetivo das ações de <i>marketing</i> .	Um ativo relacional que deve ser cultivado e alavancado.
Input para as análises de <i>marketing</i>	Entendimento do mercado e da organização.	Consequências financeiras das decisões de <i>marketing</i> .
Concepção dos ativos	Específico primeiramente à organização.	Resultado do conjunto entre a organização e o ambiente.
Participantes internos das decisões de <i>marketing</i>	Principalmente profissionais de <i>marketing</i> . Outros, somente se necessário.	Todos os gerentes que sejam relevantes, independentemente de sua função ou posição.
<i>Stakeholders</i> (externos) do <i>marketing</i>	Clientes, concorrentes, canais, reguladores.	Shareholders e potenciais investidores.
O que é medido?	Resultados de produto-mercado. Questões relacionadas a clientes, canais e concorrentes.	Resultados financeiros. Ativos baseados em mercado.
Métricas operacionais	Volume de vendas, <i>market-share</i> , satisfação dos clientes, retorno sobre as vendas, ativos e patrimônio.	Valor presente líquido dos fluxos de caixa futuros e valor para o acionista.

Fonte: Srivastava, Shervani e Fahey (1998).

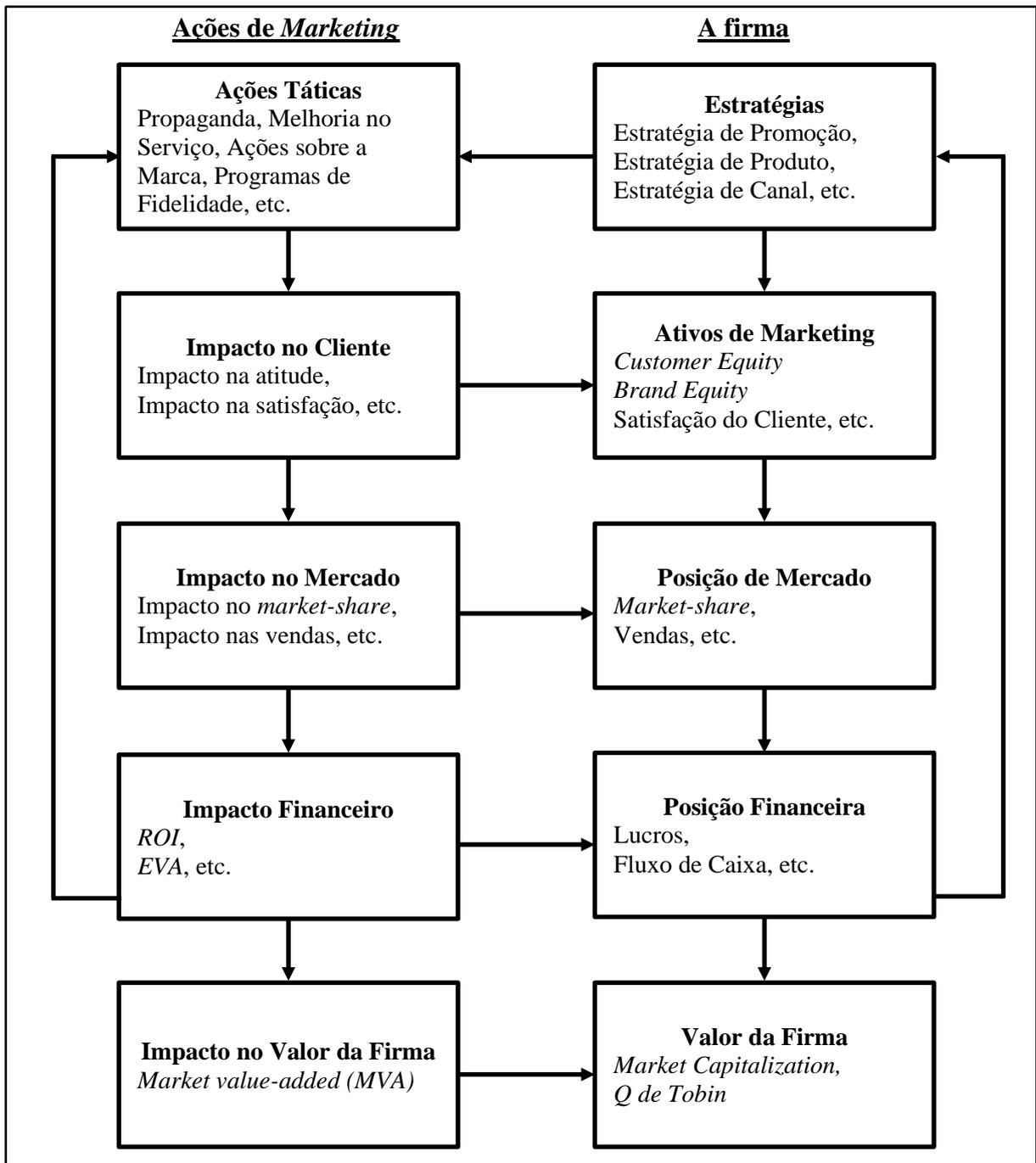
Sendo assim, durante a última década, mensurar o desempenho de *marketing* tornou-se preocupação de diversos autores da área. Rust *et al.* (2004) propõem um modelo para a mensuração da produtividade em *marketing* e sustentam que é possível mostrar como os gastos de *marketing* estão ligados ao valor do acionista. Para esses autores, quando se fala em investimento de *marketing*, é preciso identificar os ativos de *marketing* nos quais os

investimentos estão sendo aplicados, bem como compreender como esses ativos contribuem para os lucros no curto prazo e geram potencial para o crescimento e para lucros sustentáveis no longo prazo. O modelo proposto (ver Figura 1), chamado de Cadeia da Produtividade de *marketing*, busca estabelecer uma sequência lógica pela qual a produtividade de *marketing* pode ser acessada. O processo inicia-se no quadro superior direito, quando as estratégias de *marketing* envolvendo promoção, produto, canal de distribuição, entre outros, são estabelecidas. Essas estratégias devem ser implementadas por meio de ações táticas, como campanhas publicitárias, melhorias nos serviços prestados, programas de fidelidade, entre outras. Tais ações, se bem realizadas e regidas por estratégias eficazes, gerarão impacto sobre o cliente, que, por sua vez, melhorará sua atitude em relação à marca; sua satisfação com a oferta será influenciada; e a lealdade do cliente perante a marca será também influenciada. No nível da empresa, essas métricas sobre os clientes podem ser agregadas no que é chamado por Rust *et al.* (2004) de ativos de *marketing* que podem ser mensurados através de indicadores como *Brand Equity*, *Customer Equity* e Satisfação dos Clientes.

O comportamento do cliente, então, influencia o mercado, gerando mudanças no *market-share* e nas vendas, que, por sua vez, gerará maiores lucros e fluxos de caixa. Em um ponto qualquer no tempo, é importante destacar que as ações táticas de *marketing* poderão ter gerado mudanças no comportamento dos clientes, mas esse novo comportamento pode ainda não ter influenciado os lucros e os fluxos de caixa futuros da empresa. Sendo assim, é importante destacar que os ativos de *marketing* agem também como uma reserva de fluxos de caixa futuros que se acumularam a partir das atividades de *marketing* realizadas, mas ainda não se traduziram em receitas para a empresa. Por fim, uma vez mensurado o incremento sobre os ativos de *marketing* gerados pelas ações de *marketing*, é possível buscar relacionar as alterações no valor dos ativos de *marketing* com as alterações no valor da firma, calculado através de métricas como *Market Capitalization* e Q de Tobin.

Compartilhando da mesma linha de pensamento, Sheth e Sisodia (2001) afirmam que os profissionais de *marketing* precisam rever suas métricas utilizadas, pois uma mensuração da *performance* de *marketing* com base apenas em *market-share* não é suficiente: a questão correta deve ser como o *marketing* afeta o *market capitalization* da empresa. De forma mais ampla, o desafio da função de *marketing* é documentar como as atividades de *marketing* podem ser compreensivamente mensuradas em termos de suas contribuições para a *performance* financeira da companhia (SHETH e SISODIA, 2002).

Figura 1 – Cadeia da Produtividade de Marketing (RUST et al., 2004)



Fonte: Rust et al. (2004).

Gök e Hacıoglu (2010), por sua vez, consideram a produtividade de *marketing* e a gestão de sua *performance* um dos papéis emergentes do *marketing*. Para esses autores, a pressão financeira existente em diferentes indústrias fez com que o foco passasse a ser em métricas e retornos de curto prazo. Uma vez que os gastos de *marketing* são investimentos de mercado com retorno financeiro no longo prazo, a alta gerência tem geralmente considerado os gastos de *marketing* como um custo de curto prazo sem efeito financeiro observado.

Assim, a função de *marketing* tem suscitado cada vez menos interesse entre os diretores das organizações e tem gradativamente recebido menor suporte orçamentário. Em contraponto a essa realidade, Gök e Hacıoglu (2010) sugerem que, se o retorno de *marketing* pudesse ser mensurado, o papel do *marketing* dentro das organizações seria significativamente elevado.

Agora, caso isso não ocorra, os autores entendem que o *marketing* continuará a ser marginalizado, opinião que também é compartilhada por autores como Webster, Malter e Ganesan (2005) e Baker e Holt (2004). Além disso, Stewart (2009) ainda cita que, sem uma capacidade de mensurar seu desempenho, os profissionais de *marketing* se tornarão meros táticos conduzindo atividades bastante específicas que pessoas de outras áreas terão planejado.

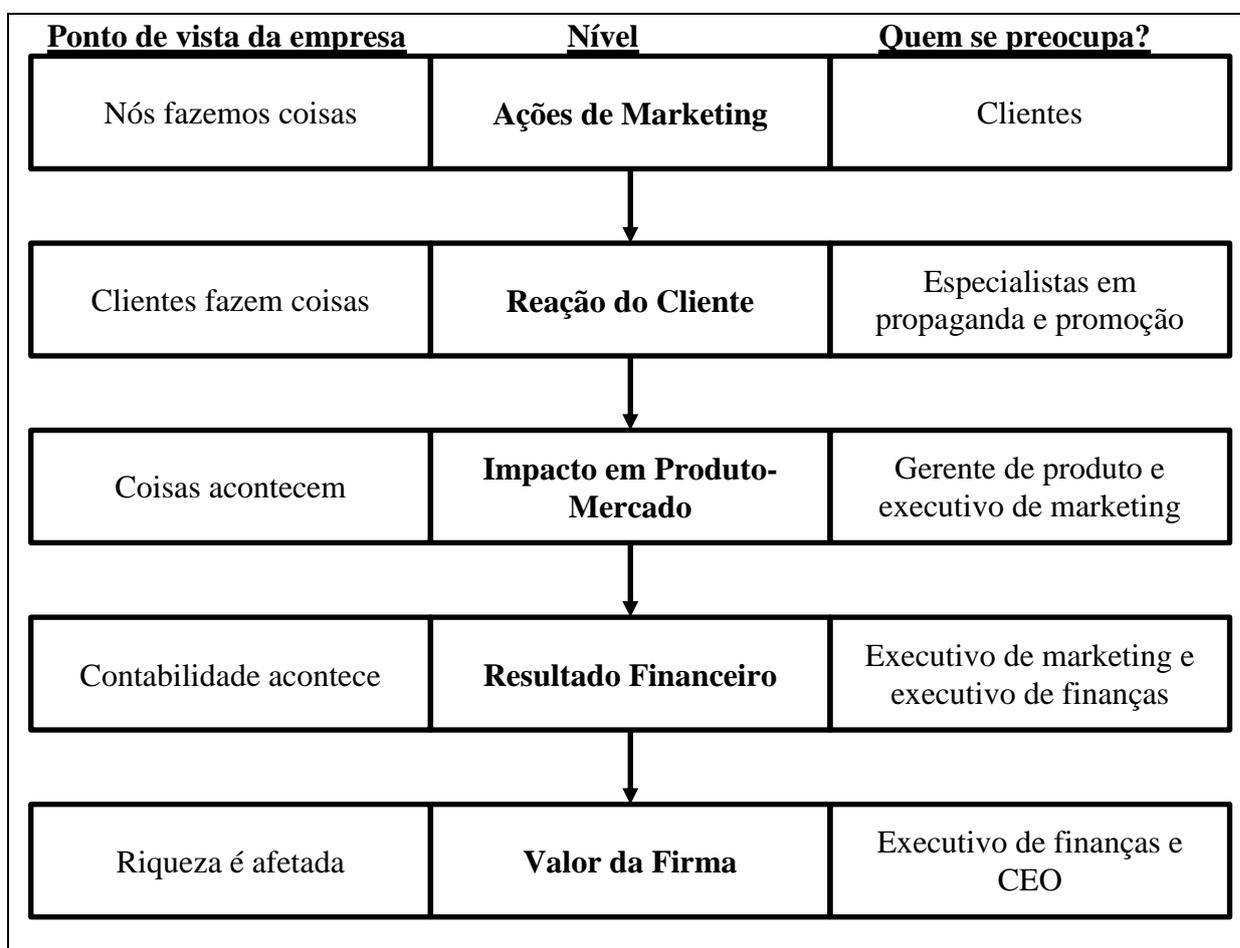
Lehmann (2004) é outro autor que também considera necessário o *marketing* ser responsável pela mensuração de sua produtividade. Segundo ele, nas décadas de 80 e 90, o antigo padrão de igualdade entre as funções de uma organização foi-se modificando até o estabelecimento da hegemonia de uma única função: finanças. Se essa realidade está certa ou não está fora do escopo de análise desse autor, ele simplesmente assume que esse é o cenário atual. Diante disso, é citado que o *marketing* utiliza-se apenas de métricas que supervalorizam resultados de curto prazo, como *ROI*. Isso não significa que essas métricas sejam ruins, porém não são suficientes para acessarem a produtividade de *marketing*. Lehmann (2004) argumenta que, para acessar a *performance*, os profissionais de *marketing* também precisam fazer uso de métricas financeiras e do valor dos ativos de *marketing*, que possuem valor de longo prazo, como *Brand Equity* e *Customer Equity*.

Colocado de outra forma, para Lehmann (2004), se o *marketing* quer um lugar na mesa em decisões importantes do negócio, é preciso relacionar seu desempenho à *performance* financeira. Caso contrário, focando apenas em métricas com as quais se sente mais confortável (atitudes, vendas, etc.), o *marketing* vai continuar a perder espaço para outras áreas. Esse autor também apresenta uma cadeia de produtividade de *marketing* (ver Figura 2), que, em essência, é bastante similar àquela desenvolvida por Rust *et al.* (2004) (ver Figura 1). Diante dessa cadeia, a tarefa que o *marketing* precisa realizar é conseguir ligar as ações de *marketing* e seus efeitos junto aos consumidores à *performance* financeira e ao valor da firma, pois é dessa forma que o CEO valorizará as ações realizadas pelo departamento de *marketing* e esse, por sua vez, poderá recuperar o espaço perdido.

Além dos autores citados, diversos outros compartilham da mesma opinião: o *marketing* precisa ser responsável por comprovar a sua produtividade para recuperar seu

espaço junto à mesa de decisões da diretoria. Sendo assim, para atingir esse objetivo, diante das argumentações dos autores citados, a capacidade de mensurar adequadamente os ativos de *marketing* torna-se essencial, pois esses ativos são considerados os elementos capazes de traduzir em valores monetários o resultado dos diversos investimentos realizados pela área de *marketing*. Diante da mensuração desses valores, é possível o planejamento e a realização de ações mais precisas por parte dos profissionais de *marketing*. Além disso, torna-se possível vislumbrar, controlando, é claro, a influência de outras variáveis, uma comparação entre variações no valor desses ativos com variações no valor da firma ou outras variáveis que representem a riqueza do acionista.

Figura 2 - Cadeia da Produtividade de Marketing (LEHMANN, 2004)



Fonte: Lehmann (2004).

A partir dos esforços iniciados na década de 90 nesse sentido, alguns importantes avanços foram observados. Entre os ativos de *marketing* existentes, os dois que suscitaram maior interesse por parte dos acadêmicos da área foram a marca (*Brand Equity*) (AAKER, 1996a, 1996b; KELLER, 1993, 1998) e os clientes (*Customer Equity*) (GUPTA, LEHMANN

e STUART, 2004; KUMAR e SHAH, 2009; RUST, LEMON e ZEITHAML, 2004). Esses ativos foram profundamente estudados, modelos diferentes para mensurar seus valores foram propostos e esforços para relacionar o valor desses ativos com o valor gerado para o acionista também foram despendidos. No presente trabalho, questões sobre o *Brand Equity* não serão diretamente abordadas, embora existam discussões importantes na literatura de *marketing* sobre a relação entre *Brand Equity* e *Customer Equity* (LEONE *et al.*, 2006; RUST, LEMON e ZEITHAML, 2004). Isso não significa que o *Brand Equity* seja menos relevante, mas apenas que não representa o presente interesse de análise, o qual está direcionado para o *Customer Lifetime Value* e *Customer Equity*. Essas duas métricas têm despontado com uma importante ferramenta para o *marketing* cumprir sua tarefa de alocar de forma eficiente e eficaz os seus recursos e comprovar os resultados de seus investimentos.

É fácil pedir milhões de dólares para propaganda ou para melhorar o serviço ao consumidor, mas é muito mais difícil mostrar como esses investimentos afetam os lucros da empresa ou o valor para o acionista. Isso fez com que as empresas recorressem a estratégias comprovadamente de curto prazo, como promoções ou corte de custos, pois seus efeitos são facilmente quantificáveis. Como uma alternativa para essa realidade, os conceitos de *CLV* (e *CE*, de forma agregada) têm ganhado importância entre os acadêmicos e práticos. Empresas como a Harrah têm obtido um tremendo sucesso gerenciando o seu negócio a partir do *CLV* e de técnicas de banco de dados. Isso porque o *CLV* permite algumas mudanças importantes na forma de gerir negócios: (1) a empresa é forçada a ser centrada nos clientes; (2) essa ferramenta foca-se na rentabilidade de longo prazo em vez de métricas de curto prazo como participação de mercado; (3) permite que a empresa acesse o valor de cada cliente e atinja-o por meio de ofertas customizadas; (4) a melhoria em tecnologia da informação e a disponibilidade de dados em nível individual permite que a empresa faça análises detalhadas em vez de confiar em métricas agregadas baseadas em pesquisa, como a satisfação (GUPTA e LEHMANN, 2006). Nas seções seguintes, é apresentada uma revisão dos conceitos de *CLV* e *CE*, bem como os benefícios de sua utilização para as empresas e alguns dos principais modelos desenvolvidos até então para mensurá-los.

2.2 CUSTOMER LIFETIME VALUE (CLV) e CUSTOMER EQUITY (CE)

Os clientes são o principal foco dos esforços de *marketing* (GUPTA, LEHMANN e STUART, 2004) e atrair e manter os clientes com maior valor é fundamental (*the cornerstone*) para um programa de *marketing* bem-sucedido (BLATTBERG e DEIGHTON,

1996). Assim, o paradigma de *CE* reconhece os clientes como a principal fonte de fluxos de caixa atuais e futuros de uma organização (VILLANUEVA e HANSENS, 2007). Além disso, diante da crescente aceitação de que o valor financeiro das firmas depende de ativos intangíveis que não estão presentes nos tradicionais balanços contábeis, o *CE* é considerado uma boa *proxy* para o valor da firma, o que faz com que a firma esteja interessada em maximizar o valor presente líquido dos fluxos de caixa atuais e futuros deixados por seus clientes (GUPTA, LEHMANN e STUART, 2004). Sendo assim, frente ao retumbante alerta, tratado na seção anterior, que ecoa entre os profissionais da área para que o *marketing* seja capaz de mensurar os seus ativos, direcionar de forma precisa suas ações, comprovar a sua produtividade e relacioná-la com o valor gerado para o acionista; a gestão dos clientes, enquanto um ativo da firma, por meio do *CLV (CE)*, é considerada um dos caminhos para que o *marketing* consiga cumprir essas tarefas e, assim, ganhar parte do seu espaço perdido perante a mesa de decisões da diretoria.

Os primeiros estudos acadêmicos sobre o tema surgiram ainda no final da década de 80, com o trabalho de autores como Schmittlein, Morrison e Colombo (1987), que buscaram identificar o tempo em que um cliente permaneceria futuramente na base de uma firma através de um modelo no qual o comportamento de compra passado do cliente era utilizado para prever, por meio do uso da distribuição binomial negativa, a probabilidade futura de o cliente permanecer fazendo negócios com a firma, ou seja, a probabilidade de o cliente se manter 'vivo' para a empresa. Jackson (1989a, 1989b e 1989c) é outro autor dessa época que tratou do tema e conseguiu ir além, buscando determinar não somente o tempo de vida de um cliente para uma empresa, como também relacionou essa estimativa com o valor monetário gerado por esses clientes, buscando mensurar, assim, o valor vitalício de um único cliente. As raízes do *CLV* estão em diversas linhas de pesquisa, incluindo *marketing* direto, qualidade do serviço, *marketing* de relacionamento e *Brand Equity*. Os profissionais de *marketing* direto foram os primeiros a utilizar a mensuração do *CLV* como base para as estratégias de *marketing* (HOGAN, LEMON e RUST, 2002). Os primeiros modelos de *CLV* que são considerados atualmente, porém, surgiram somente a partir da segunda metade da década de 90, sendo que mesmo no início dos anos 2000, autores como Hogan, Lemon e Rust (2002) apresentaram a gestão baseada no *CE* como uma nova forma de gerenciar os clientes como um ativo de *marketing*, denotando um caráter inovador para a métrica. Blattberg e Deighton (1996) foram quem propuseram o termo *CE* e desenvolveram um modelo para a mensuração do *CE*. Esse modelo foi suficiente para evidenciar a importância de compreender o valor da

base de clientes de uma firma e como utilizar esse conhecimento para determinar os níveis ótimos de investimentos em aquisição e retenção de clientes.

Antes de abordar os benefícios de uma gestão de *marketing* capaz de utilizar-se dessas métricas para o planejamento e realização de suas ações, além das características de alguns dos principais modelos existentes de *CLV* (*CE*), faz-se necessário definir esses dois conceitos e como eles são genericamente mensurados. É prudente iniciar pela definição de *CLV* ou, em Português, Valor Vitalício do Cliente. Para Villanueva e Hanssens (2007), o *CLV* é a soma descontada a valor presente dos fluxos de caixa gerados ao longo de todo o tempo futuro em que um único cliente ou segmento de clientes mantém relações de troca com uma firma. Kumar e Shah (2009), embora tratem do mesmo ponto de vista adotado por Villanueva e Hanssens (2007), definem *CLV* com mais detalhes. Para Kumar e Shah (2009), o cálculo do *CLV* envolve a previsão dos fluxos de caixa futuros de cada cliente, incorporando em uma única equação os elementos da receita gerada pelo cliente, as despesas relacionadas a esse cliente e o comportamento de compra desse cliente. Esses fluxos de caixa futuros são descontados a uma taxa de custo de capital para, então, chegar-se ao valor presente líquido de todos os fluxos de caixa futuros esperados para o cliente em questão. Borle, Singh e Jain (2008), por sua vez, também citam, com outras palavras, que uma forma genérica para o cálculo do *CLV* é dada pela estimação do valor presente dos benefícios líquidos gerados por um cliente para a firma ao longo do tempo, sendo que os benefícios líquidos são geralmente mensurados como as receitas do cliente menos os custos para a firma para manter relacionamento com esse cliente.

Embora presente nas definições apresentadas, faz-se importante salientar, portanto, que o *CLV* é o valor vitalício de um único cliente ou de um único segmento específico de clientes. Com isso, torna-se simples a compreensão de o que é *CE*. Segundo Villanueva e Hanssens (2007), o *CE* é uma métrica que busca aferir o valor dos fluxos de caixa provindos dos relacionamentos com todos os clientes de uma firma ao longo do tempo. Tratando do mesmo conceito, Kumar (2008) indica que o *CE* pode ser calculado através da soma dos *CLVs* de todos os clientes da base de uma firma. Villanueva e Hanssens (2007), porém, ainda identificam que, na literatura sobre o tema, existem duas variações para o *CE*:

(a) *Static Customer Equity (SCE)*: O *SCE* é resultado da soma de todos os *CLVs* de um grupo de clientes ou de segmentos de clientes. Os modelos estáticos de *CE* não consideram os clientes futuros, restringindo-se apenas aos *CLVs* dos clientes atuais de uma firma (YOO e HANSSENS, 2005).

(b) Dinamic Customer Equity (DCE): O *DCE* pode ser definido como a soma descontada a valor presente dos *CEs* de grupos de atuais e futuros clientes. O *DCE*, portanto, contempla o *SCE* e ainda prevê o crescimento da base de clientes. O modelo de Gupta, Lehmann e Stuart (2004) é um exemplo de *DCE*.

Villanueva e Hanssens (2007) ponderam que muitos defensores da utilização do *CE* criticam o uso de métricas que visam a maximizar apenas o retorno de curto prazo, como o *market-share*. No entanto, para esses autores, uma vez que apenas o *SCE* seja utilizado, é possível realizar a mesma crítica ao uso do *SCE*, pois esses modelos não consideram os clientes futuros de uma organização, apenas o grupo atual de clientes. De acordo com Villanueva e Hanssens (2007), modelos de *SCE* também podem levar a decisões subótimas por parte dos gestores de *marketing*. A decisão mais adequada, portanto, é a utilização de modelos dinâmicos de *CE* (*DCE*).

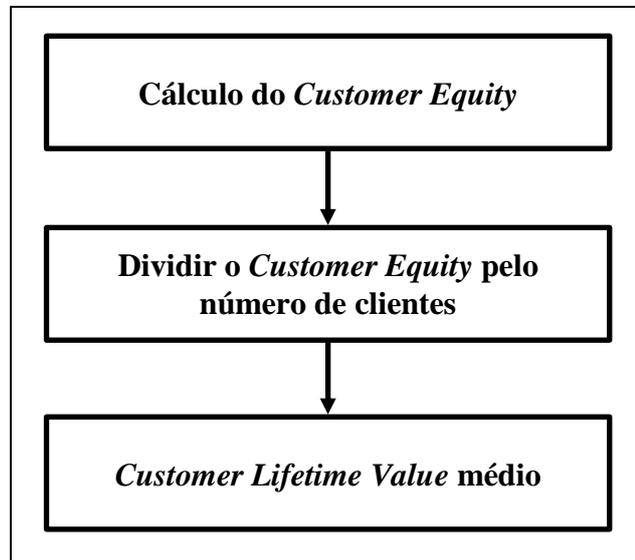
Uma vez apresentados os conceitos de *CLV*, *CE*, *SCE* e *DCE*, é necessário ainda ressaltar que existem duas maneiras de obter o *CLV*: *top-down* e *bottom-up* (KUMAR, 2008):

(a) Top-down: Por meio do modo *top-down*, o *CLV* obtido é um valor médio igual para todos os clientes. Ou seja, existem modelos que não são capazes de estimar o *CLV* de cada um dos clientes. Nesse caso, obtém-se primeiro o valor total do *CE*, e os valores dos *CLVs* são obtidos através da divisão entre o valor do *CE* e o número de clientes que a empresa possui (ver Figura 3).

O principal benefício de utilizar o modo *top-down* é a possibilidade de mensurar o *CE* sem a necessidade de dispor de informações em nível individual de todos os consumidores de uma organização. Esse modo é uma maneira simples de calcular o *CE* de uma empresa, o que, de posse de pouca informação ou de pouco tempo e recursos para viabilizar o cálculo do *CLV* de cada cliente individualmente, permite que a empresa consiga, mesmo assim, estimar o valor monetário de toda a sua base de clientes. Tais modelos são também úteis para grandes empresas, que possuem muitos clientes dispersos geograficamente, uma vez que não é necessário dispor de informações sobre todos os clientes. Além disso, em algumas situações, apenas o valor agregado da base já é suficiente para determinadas empresas. No entanto um ponto negativo evidente desse modo é que todos os clientes da empresa são considerados iguais entre si, ou seja, possuem o mesmo *CLV*. Na prática, porém, os valores dos clientes podem se diferenciar de forma significativa dentro de uma mesma base de clientes. Caso essa variabilidade exista e a necessidade seja de tratamento diferenciado para clientes distintos, Kumar (2008) recomenda a adoção de um método capaz de reconhecer as diferenças entre os

valores dos clientes. Sendo assim, o modo *top-down* possui limitada aplicação para guiar as ações de *marketing*.

Figura 3: Modo *top-down* para mensurar o *CLV*



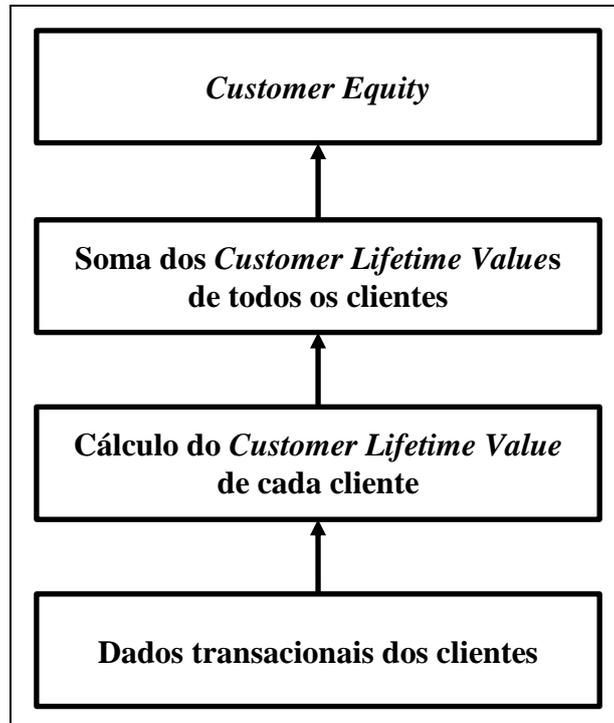
Fonte: Adaptado de Kumar (2008).

(b) *Bottom-up*: Por meio do modo *bottom-up*, primeiro o *CLV* de cada cliente da firma é estimado. A partir disso, os valores do *CLV* de cada um dos clientes são, então, somados para chegar ao valor total referente ao *CE* da firma ou de um segmento específico de clientes (ver Figura 4).

É importante destacar que um pré-requisito chave para a viabilização do cálculo do *CLV* por meio desse modo é a disponibilidade de informações em nível individual a respeito de todos clientes da base. Infelizmente, nem todas as firmas atendem a esse pré-requisito. Além disso, outro ponto negativo desse modo é o tempo e a infraestrutura de *hardware* necessária para viabilizar a mensuração do *CLV* de cada cliente, especialmente para grandes empresas, que possuem milhões de clientes. No entanto, uma vez calculado o *CLV* de forma individual, importantes *insights* podem ser alcançados sobre o comportamento individual dos clientes, as respostas a promoções, o valor individual de cada cliente, entre outros benefícios que não seriam possíveis se o modo *top-down* fosse utilizado. Além disso, o cálculo individual do *CLV* permite que segmentos de clientes sejam estabelecidos com base em previsões a respeito do valor que será gerado por cada cliente e, assim, ações diferenciadas de *marketing* podem ser definidas para cada um desses segmentos. Para Kumar (2008), buscar conhecer os clientes é necessário e, se um empresário não o faz, o seu concorrente o fará.

Sendo assim, conhecer o *CLV* de cada cliente torna-se uma importante ferramenta para alcançar esse objetivo. O cálculo do *CLV* pelo modo *bottom-up* permite que todo o potencial gerencial dessa métrica seja explorado.

Figura 4: Modo *bottom-up* para mensurar o *CLV*



Fonte: Kumar (2008).

Nos últimos anos, a ideia de gerenciar os clientes com base no *CLV* surgiu como uma das formas mais populares e eficientes de fazer negócios (KUMAR e SHAH, 2009). Para Gupta e Lehmann (2006), o *CLV* está remodelando a forma como gerenciamos para maximizar lucros. Os modelos de *CLV* são poderosas ferramentas para promover a maximização dos retornos dos investimentos de *marketing* e para guiar a alocação adequada do orçamento de *marketing* (VILLANUEVA e HANSSENS, 2007). O encanto da métrica de *CLV* está na sua capacidade de fomentar a gestão do relacionamento com o cliente por meio de intervenções adequadas de *marketing* (VILLANUEVA e HANSSENS, 2007). Kumar (2008), por sua vez, também é bastante a favor do uso do *CLV* e, assim, entende que a beleza dessa métrica está no fato de que é prospectiva (*forward-looking*), diferente de métricas tradicionais baseadas em contribuições passadas para o lucro, pois isso permite que os profissionais de *marketing* tomem ações hoje para aumentar a lucratividade futura.

Conforme já evidenciado, uma das raízes do *CLV* está no campo de pesquisa de *marketing* direto, pois o interesse desses profissionais, ao terem condições de conhecer cada um de seus clientes devido à característica personalizada dos atendimentos, é estabelecer hoje uma expectativa do que esse cliente identificado trará de benefícios líquidos no futuro. Em grandes empresas, com amplas bases de clientes, essa peculiaridade inerente ao *marketing* direto não se faz presente, o que torna difícil a empresa ter conhecimento a respeito de cada um de seus clientes e projetar as ações de *marketing* com base nessas informações. A mensuração do *CLV*, diante desse cenário, proporciona a estimação do valor vitalício de cada cliente de uma empresa, mesmo que ela possua uma ampla base de clientes, e os valores encontrados podem ser utilizados para realizar ações de *marketing* mais eficientes e eficazes (BORLE, SINGH e JAIN, 2008). Indo além, esses autores afirmam que, diante do fato de que os orçamentos de *Marketing* são limitados, tendo como base o valor esperado de cada cliente, é possível estabelecer estratégias diferentes para diferentes clientes ou segmentos de clientes, o que auxilia a firma a obter melhores retornos nos seus investimentos de *marketing*.

Berger *et al.* (2002) também entendem que a alocação adequada dos recursos *marketing* para maximizar o valor da base de clientes é o caminho certo a seguir, no entanto, para que isso seja possível, os autores sustentam que é preciso, primeiramente, construir uma base de dados a partir de uma inteligência de *marketing* que seja capaz de reter informações válidas e confiáveis a respeito do histórico de compras dos clientes, das ações de *marketing* da firma em relação aos clientes e das características demográficas dos clientes. Tendo essas informações eletronicamente organizadas, modelos estatísticos para a previsão do *CLV* e segmentação da base de clientes podem ser aplicados, propiciando, então, a adequada alocação dos recursos de *marketing*.

Diante disso, Kumar e Shah (2009) utilizaram as informações disponíveis na base de dados de uma empresa de *B2B* que fabrica *hardware* e *software* para computadores para estimar os *CLVs* dessa empresa pelo modo *bottom-up* e direcionar as atividades de *marketing* junto aos clientes. Os autores conseguiram segmentar a base de clientes a partir do *CLV*. Os resultados mostraram que 20% dos clientes eram responsáveis por 91% dos lucros e que outros 20% dos clientes apresentaram *CLV* negativo. A base de clientes foi, então, segmentada entre os clientes com alto *CLV* (20% mais lucrativos), médio *CLV* (60% intermediários) e *CLV* negativo (20% com valor negativo). A partir dessa segmentação, as ações de *marketing* da empresa, durante dez meses e sem alterar o valor do orçamento da área, seguiram a seguinte regra: (a) os recursos de *marketing* foram realocados para focar na

aquisição de clientes que atendessem ao perfil dos clientes do segmento com alto *CLV*; (b) a estratégia multicanal foi de incentivar os clientes com alto e médio *CLV* a comprarem em mais de um canal, enquanto os clientes do segmento com *CLV* negativo foram direcionados para comprar apenas pelo canal *on-line*; (c) os privilégios de compra e recompensas especiais passaram a ser concedidas com base no *CLV*. A implementação dessas estratégias teve impacto direto nos direcionadores do *CLV* (margem de contribuição, quantidade de comprada, *cross-buying*, entre outros). O resultado foi o aumento de 19,4% no *CE*, comprovando, assim, o potencial do *CLV* (*CE*) para direcionar ações eficientes e eficazes por parte do departamento de *marketing*, bem como comprovar o desempenho dessas ações.

Produtos vão e vêm, mas os clientes permanecem (RUST, LEMON e ZEITHAML, 2001). Com o auxílio de métricas focadas no cliente como *CLV* e dos benefícios gerados pela sua utilização para a gestão de *marketing*, conforme evidenciado por Kumar e Shah (2009), gerentes de *marketing* de empresas líderes de mercado estão organizando os esforços de *marketing* em torno dos clientes e não em torno de produtos. Nessas empresas, o conceito orientado ao produto trazido pelo *Brand Equity* está sendo suplantado pelo conceito focado nos clientes do *CE* (HOGAN, LEMON e RUST, 2002). O *CLV* é visto como uma métrica adequada para promover a aquisição e retenção dos clientes ‘certos’ em uma gestão de relacionamento com clientes (VENKATESAN e KUMAR, 2004), o que indica que o conceito de *CLV* está baseado na visão de que a empresa deve buscar estabelecer relacionamentos com os clientes, no entanto não com todo e qualquer cliente, mas com aqueles que geram maior valor.

Para Gupta e Lehmann (2006), o valor do cliente possui dois lados: o valor que a empresa gera para o cliente (visão tradicional de *marketing*) e o valor que o cliente gera para a empresa (visão financeira). Uma empresa, ao alocar os seus recursos, precisa considerar os dois lados. Isso significa que querer mais clientes e mais *market-share* pode não ser necessariamente o melhor para a firma. A Best Buy, por exemplo, reconheceu isso e cortou os recursos que eram destinados aos seus clientes com baixo valor. Em contrapartida, a empresa investiu nos clientes de maior valor. A razão para fazer isso é obter um maior retorno dos clientes, pois assim eles irão prover lucros correntes e futuros para a firma (medido por meio do *CLV*). Essa perspectiva não só faz com que o *marketing* seja capaz de comprovar os resultados de seus gastos (ex.: programas podem ser avaliados com base no seu impacto sobre o *CLV*), mas também sugere que esses gastos de *marketing* devem ser vistos como investimentos (como ocorre P&D, por exemplo) e não como custos.

Ao buscar, por meio do *CLV*, determinar os clientes ‘certos’, a empresa passa a discriminar segmentos distintos geradores de receitas identificando de onde vem a receita, que, tradicionalmente, é unificada em apenas uma ou duas linhas no topo da demonstração de resultado do exercício. Com essa prática, as empresas passam a poder gerenciar não somente os custos de forma segmentada, mas também as receitas, atendendo, assim ao que é requisitado por Ambler (2003), que afirma que os gerentes despendem nove vezes mais atenção para os gastos e para o cálculo de fluxos de caixa do que vislumbrando de onde esses fluxos de caixa vêm e como podem ser aumentados, ou seja, é necessário também preocupar-se com a origem das receitas e como aumentá-las.

Por fim, embora não seja o foco do presente estudo, é preciso destacar que existe ainda um forte movimento dentro da área de *marketing* para relacionar o valor dos ativos de *marketing* com o valor da firma. A partir de uma ampla revisão da literatura realizada por Srinivasan e Hanssens (2009), foi possível identificar que pesquisas anteriores buscaram relacionar ao valor da firma variáveis de *marketing* como satisfação (ANDERSON, FORNELL e MAZVANCHERYL, 2004; FORNELL *et al.*, 2006) *Brand Equity* (MIZIK e JACOBSON, 2008; RAO, AGARWAL e DAHLOFF, 2004) propaganda (JOSHI e HANSENS, 2009), anúncio de novos produtos (SORESCU, SHANKAR e KUSHWAHA, 2007), promoção de preço (PAUWELS *et al.*, 2004), entre outras. Diante disso, é natural que o *CE* esteja também entre essas variáveis, especialmente pelo fato de seu cálculo estar baseado na expectativa de fluxos de caixa futuros. Sendo assim, autores como Gupta, Lehmann e Stuart (2004), Kumar e Shah (2009) e Gupta e Lehmann (2006) tratam da relação entre *CE* e o valor da firma. Esse movimento é motivado pela necessidade de o *marketing* recuperar seu espaço estratégico dentro das organizações. Gupta e Lehmann (2006) sugerem que o *CLV* (*CE*), embora comprovadamente eficiente para tornar o *marketing* mais capaz de comprovar os resultados de seus investimentos, ainda permanece sendo considerado tático sob a visão dos diretores e do mercado financeiro. Diante disso, eles argumentam que esses atores deveriam considerá-lo estratégico e, para que isso seja possível, mostram que o *CE* pode e tem sido ligado ao valor de firma.

2.3 MODELOS DE ESTIMAÇÃO DO *CUSTOMER LIFETIME VALUE* (*CUSTOMER EQUITY*)

Os conceitos de *CLV* e *CE* são relativamente simples, de forma que a complexidade do tema está relacionada com o método utilizado para estimar esses valores, as variáveis

utilizadas para determiná-los, bem como construir e organizar a base de dados sobre os clientes para que seja possível reunir as informações necessárias para a estimação. Diante disso, a maior parte da literatura sobre o tema diz respeito aos esforços para construir modelos melhores para estimar os valores futuros dessas variáveis (GUPTA e LEHMANN, 2006). Nas subseções a seguir, alguns dos principais modelos de *CLV* (*CE*) frequentemente citados na literatura serão brevemente analisados com o objetivo de demonstrar a vasta gama de métodos existentes para acessar um mesmo valor e também para fornecer subsídios de o porquê os modelos de Kumar *et al.* (2008) e de Kumar e Shah (2009) foram tão importantes para o estudo conduzido no presente trabalho.

2.3.1 Jain e Singh (2002)

Embora esses autores não tenham desenvolvido nenhum modelo original para a mensuração do *CLV*, eles merecem destaque, pois, à luz do conhecimento sobre o tema existente na época, eles fazem uma revisão dos principais modelos existentes. Entre esses, é importante apresentar o modelo estrutural básico de *CLV*, pois evidencia a origem do cálculo do *CLV*, cuja essência ainda permanece presente nos modelos mais sofisticados existentes nos dias atuais, embora evoluções tenham sido propostas. O modelo a seguir descreve de forma simples o princípio por trás do cálculo do *CLV* que é de estimar os fluxos de caixa futuros que serão gerados por um cliente:

$$CLV = \sum_{i=1}^n \frac{(R_i - C_i)}{(1 + d)^{i-0,5}} \quad \text{Equação 1}$$

Nessa equação, i = período do fluxo de caixa; R_i = receita do consumidor no período i ; C_i = custo total para gerar a receita R_i no período i ; n = número total de períodos do tempo vitalício do cliente analisado. Esse modelo, obviamente, desconsidera futuras aquisições de clientes, desconsidera a natureza estocástica do processo de compra, por ser um modelo determinístico, entre outras simplificações. Sua inclusão entre os modelos de *CLV* se dá devido ao fato de que é uma das formas mais cruas de mensurar o *CLV*. Para encontrar o *CE*, diante dessa ótica, seria necessário somar os valores dos *CLVs* encontrados para cada um dos clientes.

Usar o modelo estrutural básico para estimar o *CLV* significa tomar o tempo vitalício esperado para o cliente, o tempo médio entre compras, o volume monetário das compras passadas e os custos rateados por clientes e usar esses valores médios para prever o valor presente dos fluxos de caixa futuros (SINGH e JAIN, 2010). Esse tipo de modelo heurístico seria um método útil para calcular *CLV* na falta qualquer outro modelo melhor (BORLE, SINGH e JAIN, 2008).

2.3.2 Berger e Nasr (1998)

Nesse estudo, são apresentadas e discutidas variações do modelo estrutural básico para mensurar o *CLV*. O foco dos autores está sobre a determinação da margem de contribuição líquida de um cliente, uma vez que tenha sido adquirido. Diante dessa abordagem, os custos de aquisição, que naturalmente são importantes para uma firma, não são considerados. Os modelos tratados por esses autores, portanto, são estáticos. O modelo a seguir é o mais simples daqueles sugeridos pelos autores:

$$CLV = \left\{ GC * \sum_{i=1}^n \left[\frac{(r^i)}{(1+d)^i} \right] \right\} - \left\{ M * \sum_{i=1}^n \left[\frac{(r^{i-1})}{(1+d)^{i-0,5}} \right] \right\} \quad \text{Equação 2}$$

GC = é a margem de contribuição anual esperada por cliente (calculada por meio do valor das vendas menos os seus custos); M = custos promocionais (custos de *marketing*) por consumidor por ano; n = número total de períodos, em anos, pelo qual os fluxos de caixa precisam ser projetados; r = taxa de retenção anual (a proporção esperada de clientes que continuarão comprando da empresa no ano seguinte); d = taxa de desconto apropriada para a empresa; i = período do fluxo de caixa.

Esse modelo apresenta alguns avanços em relação ao modelo estrutural básico, pois separa os custos de venda dos custos de *marketing* e considera a possibilidade de deserção dos clientes da base ao considerar a taxa de retenção no cálculo do *CLV*. Esse modelo, porém, requer algumas premissas que desconsideram a natureza estocástica do processo de compra: o custo de *marketing* (M) permanece constante, a taxa de retenção (r) permanece constante e a margem de contribuição (GC) também permanece constante.

O último modelo proposto pelos autores no artigo busca compreender uma maior variabilidade no cálculo do *CLV* por meio da mensuração da quantidade de clientes que

estarão presentes na base a cada período do fluxo de caixa futuro e da tentativa de imputar uma probabilidade de compra por parte do cliente, apesar de a forma de cálculo dessa probabilidade ainda ser determinística. Embora ainda incipiente, essa tentativa se mostrou bastante válida, uma vez que modelos futuros propuseram a estimação da probabilidade de compra considerando sua natureza estocástica, o que trouxe grande evolução para o cálculo do *CLV* (ex.: KUMAR E SHAH, 2009).

2.3.3 Blattberg e Deighton (1996)

Enquanto os modelos de Berger e Nasr (1998) aplicam-se somente para clientes já adquiridos pela empresa (pessoas ou empresas que já compraram da empresa focal), o modelo de Blattberg e Deighton (1996) aplica-se especificamente para prospecções (pessoas ou empresas que ainda comprarão da empresa focal) (PFEIFER e CARRAWAY, 2000). O principal objetivo de Blattberg e Deighton (1996) é auxiliar os gestores a determinar o ponto ótimo entre custos de aquisição e custos de retenção de clientes, ou seja, quanto de recursos devem ser alocados em cada uma dessas duas frentes para maximizar a rentabilidade. Esse é um modelo de que não considera a variabilidade entre os *CLVs* de cada cliente, ou seja, o cálculo do modelo gera um valor agregado para toda a base de clientes da empresa (*CE*). Diante disso, o ponto ótimo entre custos de aquisição e custos de retenção de clientes é obtido quando o *CE* é maximizado. A expressão a seguir resume a relação encontrada pelos autores:

$$CE = a \cdot m - A + a \left(m - \frac{R}{r} \right) \left[\frac{r'}{(1 + r')} \right] \quad \text{Equação 3}$$

Nessa expressão, *A* = custo de aquisição; *R* = custo de retenção; *a* = taxa de aquisição para o nível de custo *A*; *r* = taxa de retenção para o nível de custo *R*; *m* = margem de contribuição da firma; $r' = r/(1+d)$, sendo *d* = taxa de desconto. A partir dessa expressão, fica claro que o problema principal é o cálculo dos custos de aquisição e retenção ótimos, no entanto pode ser também utilizado, diante de um dado nível de gastos, para calcular o *CE*. A definição dos valores de entrada de cada uma dessas variáveis se dá de forma determinística, a partir de expectativas informadas pelo gerente/diretor da empresa.

Para obter o valor do *CLV* a partir de um modelo agregado como esse, é preciso dividir o valor do *CE* pelo número de clientes para obter um valor médio para o *CLV*. É

evidente que, nesse caso, os problemas citados por Kumar (2008), e destacados anteriormente, a respeito do modo *top-down* de obtenção do *CLV* estão presentes. No entanto, uma vez que a empresa esteja interessada em apenas gerenciar os custos de aquisição e retenção ou que não possua recursos suficientes para promover uma mensuração desagregada do *CLV*, esta se torna uma opção interessante.

2.3.4 Rust, Lemon e Zeithaml (2004)

O modelo proposto por esses autores utiliza Cadeias de Markov, que há muitos anos vem sendo utilizadas em *marketing* para modelar a troca de marcas por parte dos clientes. A flexibilidade e habilidade para modelar a competição proporcionada por esse método permite que ele seja utilizado para modelar também o *CLV*. Sendo assim, quando dados sobre múltiplas empresas que competem por um grupo de clientes estão disponíveis, esse modelo é uma opção muito boa para estimar o *CLV* de uma empresa específica. A flexibilidade desse modelo permite, ainda, a estimação do impacto de diversos direcionadores (*drivers*) no valor vitalício dos clientes da empresa focal e de seus concorrentes. Além disso, o modelo de Rust, Lemon e Zeithaml (2004) é particularmente importante devido à forma pela qual as informações necessárias para a montagem da Matriz de Markov e avaliação dos direcionadores do *CLV* são obtidas, pois isso é realizado por meio de pesquisa (*survey*) *cross-sectional* com uma amostra dos clientes de uma empresa (SINGH e JAIN, 2010). Isso faz com que a mensuração do *CLV* seja possível mesmo para empresas com um número muito elevado de clientes e, principalmente, quando a relação entre cliente e empresa é não contratual.

Os direcionadores do *CLV* e seus subcondutores são: valor para o cliente (qualidade preço e conveniência), *Brand Equity* (lembrança de marca, atitudes em relação à marca e percepção de ética da marca) e relacionamento ou *Retention Equity* (programas de lealdade, tratamentos especiais, criação de conhecimento sobre a empresa por parte do cliente e criação de comunidade). Os direcionadores são mensurados a partir de escalas Likert de 5 pontos e, uma vez que os impactos desses direcionadores sobre o *CE* também são mensurados, é criada uma importante oportunidade para os gestores tomarem conhecimento de quais direcionadores mais impactam no *CE* e, assim, priorizar as ações futuras de *marketing* que tratem daqueles direcionadores com maior impacto sobre o *CE*. Essas ações, porém, devido ao fato de os *CLVs* obtidos serem iguais para todos os clientes, podem ser tomadas apenas de

forma agregada, direcionando-se os recursos de *marketing* para os indicadores de cada um dos *drivers* que mais impactam no *CE*.

O cálculo do *CLV* de cada um dos clientes participantes da *survey* é realizado a partir da seguinte equação:

$$CLV_{ij} = \sum_{t=1}^{T_{ij}} \frac{1}{(1+d)^t} * v_{ijt} * \pi_{ijt} * B_{ijt} \quad \text{Equação 4}$$

Diante disso, o volume de compra (*v*) multiplicado pela margem de contribuição esperada (π) e pela probabilidade de compra (*B*) trazido a valor presente pela taxa de desconto (*d*) representa o *CLV* do cliente *i* para a firma *j*.

Nessa equação, a variável que precisa ser calculada é a probabilidade de compra. Usando informações individuais sobre cada cliente da amostra e combinando isso com os dados de intenções de compra desses clientes, cada matriz de troca entre as empresas constituintes do mercado é montada. Cada cliente *i* tem, então, uma matriz de troca *JxJ*, na qual *J* é o número de marcas constituintes do mercado consideradas. A probabilidade individual de escolha (compra) na matriz de troca (ex.: probabilidade que um cliente *i* escolha uma marca *k*, dado que a marca *j* foi a marca escolhida na última compra) é calculada usando o modelo logit multinomial.

Uma vez obtidos os valores dos *CLVs* para cada um dos clientes amostrados, o *CE* é obtido por meio da seguinte equação:

$$CE = \text{média (CLV)} * POP \quad \text{Equação 5}$$

Dados os valores dos *CLVs* da amostra obtida, calcula-se o *CLV* médio e, supondo-se que a amostra é representativa da população de clientes, multiplica-se esse valor médio pelo número total de clientes da empresa focal (*POP*) e obtém-se o valor do *CE*. Torna-se evidente, assim, que a desvantagem desse modelo é que novamente o valor do *CLV* obtido é um valor médio e igual para todos os clientes da base, o que impossibilita os gestores de compreenderem quais são os clientes de maior ou menor valor e de realizarem ações personalizadas entre os clientes ou entre segmentos de clientes.

2.3.5 Gupta, Lehmann e Stuart (2004)

O modelo desses autores é particularmente especial, pois é desenvolvido para que possa ser calculado a partir de informações públicas divulgadas pelas empresas. Sendo assim, esse modelo não requer a disponibilidade de grandes bases de dados. Além disso, o modelo é conceitualmente simples e é maior a facilidade de cálculo do *CE* e do *CLV* médio de cada cliente (modo *top-down*). Para tanto, é feito uso de uma variável capaz de captar a probabilidade de deserção dos clientes da base, que é a taxa de retenção. Essa variável, que já havia sido utilizada em modelos anteriores como o de Blattberg e Deighton (1996), assume papel fundamental para o modelo Gupta, Lehmann e Stuart (2004) e é considerada constante, embora os próprios autores reconheçam que essa é uma simplificação que pode não ocorrer na realidade. Nessa caso, a fórmula básica para o *CLV* segundo essa abordagem é dada por:

$$CLV = \sum_{t=0}^{\infty} m_t \frac{r^t}{(1+i)^t} \quad \text{Equação 6}$$

Onde m = margem de contribuição média (constante), r = taxa de retenção (constante) e i = taxa de desconto. A partir dessa fórmula mais elementar, é possível verificar que o tempo considerado para o cálculo do valor presente do fluxo de caixa provido pelos clientes é considerado infinito. Segundo os autores, modelos com horizontes infinitos são mais simples de serem estimados.

A partir dessa fórmula fundamental, é construído o modelo de *CE*, considerando toda a base de clientes. Outra característica importante é que os autores incluem também as aquisições futuras de clientes na versão mais sofisticada do modelo, tornando-o dinâmico. Sendo assim, a fórmula para a mensuração do *CE* dinâmico de maneira discreta segue:

$$DCE = \sum_{k=0}^{\infty} \frac{n_k}{(1+i)^k} \sum_{t=k}^{\infty} m_{t-k} \frac{r^{t-k}}{(1+i)^{t-k}} - \sum_{k=0}^{\infty} \frac{n_k c_k}{(1+i)^k} \quad \text{Equação 7}$$

Onde n = número de clientes, k = é o grupo de cliente existente em um período t (a cada período um novo grupo de clientes é adquirido), m = margem de contribuição média por cliente, t = tempo e c = custo de aquisição médio por cliente. Embora seja mais simples conceituar o modelo em termos discretos, na realidade, a aquisição e retenção de clientes é um

processo contínuo. Sendo assim, os autores utilizam uma versão contínua do modelo para calcular o *CE*:

$$DCE = \int_{k=0}^{\infty} \int_{t=k}^{\infty} n_k m_{t-k} e^{-ik} e^{-\left(\frac{1+i-r}{r}\right)(t-k)} dt dk - \int_{k=0}^{\infty} n_k c_k e^{-ik} dk \quad \text{Equação 8}$$

Embora o modelo de Gupta, Lehmann e Stuart (2004) possua as vantagens já referidas, é necessário destacar também suas limitações. O modelo dificilmente seria aplicável para um cenário não contratual, pois a taxa de retenção é difícil de ser observada em empresas com relação não contratual, uma vez que fica difícil identificar quando o cliente está ativo ou não. Além disso, os custos de aquisição são calculados a partir da divisão dos custos de *marketing* pelo número de novos clientes, o que é uma suposição bastante forte que desconsidera o caso de empresas já maduras que investem pesado em retenção de clientes (VILLANUEVA e HANSSENS, 2007). Por fim, o modelo segue o modo *top-down* para o cálculo do *CLV*, o que impossibilita a consideração do *CLV* de forma individual para a alocação dos recursos de *marketing*, pois o valor vitalício é igual para todos os clientes.

2.3.6 Schmittlein, Morrison e Colombo (1987) / Schmittlein e Peterson (1994)

O modelo de Schmittlein, Morrison e Colombo (1987), conhecido como Pareto/*NBD* (Pareto/*Negative Binomial Distribution*), busca mensurar, com base no comportamento de compra passado dos clientes, a probabilidade de um cliente se manter ‘vivo’ em um relacionamento com uma firma. De acordo com Singh e Jain (2010), o modelo Pareto/*NBD* pode ser usado para estimar (a) o número de clientes atualmente ativos; (b) como esse número muda ao longo do tempo; (c) a probabilidade de que um cliente específico esteja ativo; (d) por quanto tempo é provável que um cliente se mantenha ativo; (e) o número esperado de transações de um cliente durante um intervalo futuro de tempo.

Para atingir tais objetivos, porém, os autores estabelecem algumas suposições para que o seu modelo seja aplicável, como:

- (a) Enquanto um cliente está ativo em um relacionamento com uma firma, ele realiza transações com essa firma que são aleatoriamente distribuídas no tempo com uma taxa λ específica para cada cliente. A taxa de transação λ segue uma distribuição Gamma;

- (b) O número X de transações realizadas por um cliente no período de tempo t segue distribuição de Poisson;
- (c) Um cliente possui um tempo de vida não observado representado por τ , que é uma variável aleatória exponencial.
- (d) A taxa de transação λ , e a taxa de deserção μ variam independentemente entre os clientes.

Assumindo essas suposições, Schmittlein e Peterson (1994) validam o modelo de Schmittlein, Morrison e Colombo (1987) e o estendem incorporando o valor médio em dólares gasto pela base de clientes, o que, cruzado com a probabilidade de cada cliente se manter vivo na base, seria uma boa estimativa do valor que cada cliente traria futuramente. Além disso, uma nova suposição é estabelecida: o valor médio gasto entre os clientes é independente da distribuição da taxa de transação λ e da taxa de deserção μ .

Sendo assim, embora bastante auspicioso e inspirador, o modelo Pareto/*NBD* está baseado em suposições fortes a respeito do comportamento de compra dos clientes de uma firma, o que pode não ser aplicável em muitas situações (BORLE, SINGH e JAIN, 2008). Além disso, ao utilizarem apenas o valor médio gasto por todos os clientes, o modelo negligencia informações que podem ser importantes sobre o comportamento individual dos clientes.

2.3.7 Fader, Hardie e Lee (2005)

O modelo desses autores é uma alternativa mais simples para estimar o modelo Pareto/*NBD* e é conhecido como Beta-Geométrico/*NBD* (BG/*NBD*) (Singh e Jain, 2010). Para tanto, o modelo possui suposições levemente diferentes sobre o comportamento de compra dos clientes de uma firma que fazem com que seja mais simples sua implementação, porém mantendo resultados similares.

As principais suposições desse modelo são:

- (a) Enquanto um cliente está ativo em um relacionamento com uma firma, ele realiza transações com essa firma que são aleatoriamente distribuídas no tempo com uma taxa λ específica para cada cliente. A taxa de transação λ segue uma distribuição Gamma;
- (b) O número X de transações realizadas por um cliente no período de tempo t segue distribuição de Poisson;

- (c) Depois de cada transação, o cliente deserta (se torna inativo) com probabilidade p . Assim, a deserção dos clientes é distribuída entre as transações como uma distribuição geométrica.
- (d) p segue uma distribuição beta.
- (e) λ e p variam de forma independente entre os clientes.

Em seu artigo, Fader, Hardie e Lee (2005) comparam o seu modelo com o modelo Pareto/*NBD* e confirmam que ambos são precisos tanto no nível individual do cliente quanto no nível agregado de todos os clientes analisados. Apesar de esse modelo ser uma alternativa plausível ao modelo de Pareto/*NBD*, as mesmas críticas ao modelo de Pareto/*NBD* se aplicam também ao modelo BG/*NBD*.

2.3.8 Modelos *RFM* (Em Inglês: *Recency, Frequency and Monetary*)

Existem diversos autores que trabalham com o tradicional modelo *RFM* (BORLE, SINGH e JAIN 2008; DONKERS, VERHOEF e JONG, 2007; MALTHOUSE e BLATTBERG, 2005; REINARTZ e KUMAR, 2003). Os modelos *RFM* utilizam informações sobre o comportamento de compra passado dos clientes em relação a três dimensões (recência da compra passada, frequência das compras passadas e o valor monetário das compras passadas) para dar notas aos clientes. Diante dessas informações, modelos de regressão podem usar, então, essas entre outras variáveis para criar uma nota para cada cliente. A nota está relacionada ao potencial de compra esperado de cada cliente e, assim, pode ser considerada outra métrica de valor do cliente para a empresa. As notas geralmente servem como resultados similares aos valores de *CLV*.

Reinartz e Kumar (2003) citam que os modelos de *RFM* são comumente utilizados por empresas de *marketing* direto. No entanto, para esses autores, os modelos de *RFM* possuem limitações. Os clientes são tratados como uma única população e a eles são conferidas notas baseadas na mesma regra de decisão, embora eles pertençam a dois segmentos com comportamentos de compra diferentes. Diante disso, recursos são desperdiçados em correspondências que nunca levarão ao comportamento de compra, simplesmente porque o padrão de atividade histórica de certo segmento não é levado em consideração. Quando segmentos de clientes apresentam diferentes padrões de atividade de compra, torna-se uma importante decisão gerencial, por exemplo, deixar de enviar correspondências para clientes bastante ativos, porém com vida muito curta como comprador.

2.3.9 Borle, Singh e Jain (2008)

O modelo desses autores é particularmente importante, pois esses autores utilizam um modelo bayesiano hierárquico para a estimação das variáveis dependentes para o cálculo do *CLV*. Essas variáveis são: probabilidade de o cliente deixar de ser cliente, valor monetário de cada compra e o intervalo de tempo entre cada compra. Para informações sobre as distribuições *a priori* estabelecidas para cada uma dessas variáveis, bem como as equações utilizadas para estimar cada uma dessas variáveis, obter acesso a Borle, Singh e Jain (2008).

O que se faz importante é a exposição da forma pela qual o autor estimou o *CLV*, bem como a maneira como, por meio de estatística bayesiana, a empresa pode, a cada nova compra de um cliente, calcular novamente o valor vitalício desse indivíduo de forma atualizada, devido aos novos dados existentes. Os autores, com base em uma amostra específica de 1000 clientes, definem as funções *a priori* dos parâmetros necessários para calcular cada uma dessas três variáveis utilizadas para definir o *CLV*. De posse dessas *prióris*-padrão, quando cliente em específico compra pela primeira vez, uma simulação é feita, com base na probabilidade de deserção padrão, para identificar, de forma simulada, quantas compras ele faria até que deserte. Com base em um fluxo de caixa, incluindo o número de ocasiões de compra e o valor monetário de cada compra, que é formado a partir dessa simulação, o *CLV* é mensurado por meio do cálculo do valor presente desse fluxo de caixa. Quando uma nova compra é feita pelo mesmo cliente, o sistema passa a ter uma nova informação que é específica desse cliente (valor monetário da compra, intervalo de tempo desde a primeira compra e a certeza de que ele continua vivo após a compra anterior). Sendo assim, a simulação anteriormente explicada é realizada novamente, e o *CLV* é outra vez calculado, porém, nesse momento, as informações específicas desse cliente, adquiridas após a primeira compra, são utilizadas junto com as informações das *prióris*-padrão, tornando, assim, atualizada e mais precisa a previsão. Esse exemplo ilustra a capacidade de atualização dinâmica que pode ocorrer no cálculo do *CLV* quando se utiliza um método bayesiano (esse encadeamento de uma compra após outra não foi, porém, explorado por Kumar e Shah (2009) em seu modelo).

Apesar desse benefício, o modelo dos autores não considerou os gastos que a empresa teve para entregar a oferta ao cliente, nem quanto é gasto para realizar eventuais contatos de *marketing* com esse cliente. Dessa forma, o valor vitalício do cliente torna-se superestimado.

Além disso, o modelo, segundo considerações dos próprios autores, possui limitações, pois é melhor aplicável quando a empresa tem condições de saber exatamente quando o cliente torna-se cliente e quando ele deixa de ser um cliente, o que é um privilégio do qual muitas empresas não compartilham.

2.3.10 Venkatesan, Kumar e Bohling (2007)

O modelo utilizado por esses autores apresenta uma forma de cálculo do *CLV* que é bastante semelhante à forma utilizada tanto por Kumar *et al.* (2008) e Kumar e Shah (2009), que serão apresentadas a seguir, porém as variáveis estimadas são distintas. Para realizar tal estimação, todos esses três modelos citados utilizam-se de estatística bayesiana. O modelo de Venkatesan, Kumar e Bohling (2007) é construído sobre uma abordagem “*always-a-share*” para a relação entre cliente e empresa, o que, segundo os autores, é mais apropriado para o contexto de relação não contratual existente na empresa que eles utilizaram para testar o modelo. Rust, Lemon e Zeithaml (2004) indicam que assumir uma relação “*always-a-share*” significa que os clientes podem não dar para a firma focal todos os seus negócios, ou seja, assume-se que os clientes compram também de concorrentes. Sendo assim, diante dessa abordagem, Venkatesan, Kumar e Bohling (2007) mensuram *CLV* através da previsão de padrões de compra dos clientes ao longo de um período razoável de tempo e não através da previsão de quando os clientes vão terminar o seu relacionamento com a firma focal. O período de tempo futuro escolhido, então, é de 36 meses, que é citado como um período de tempo adequado para realizar previsões supondo-se que a dinâmica de mercado não irá mudar.

O modelo de *CLV* de Venkatesan, Kumar e Bohling (2007) é calculado a partir da seguinte equação, que representa os fluxos de caixa futuros gerados por cada cliente da empresa:

$$CLV_i = \sum_{j=T^*+1}^{T^*+T_i} \frac{\hat{Q}_{i,j} * M}{(1+r)^{\hat{t}_{i,j}}} - \sum_{t=1}^n \frac{\sum_q C_{i,q,t} * X_{i,q,t}}{(1+r)^{t-1}} \quad \text{Equação 9}$$

Onde:

CLV_i = Valor vitalício do cliente i ;

$\hat{Q}_{i,j}$ = Quantidade de compra prevista para o cliente i na ocasião de compra j ;

M = Margem de contribuição para um único item;

r = Taxa de desconto;

$c_{i,q,t}$ = Custo unitário de *marketing* para o cliente i no canal q no ano t ;

$x_{i,q,t}$ = O número de contatos de *marketing* com o cliente i no canal q no ano t ;

$\hat{t}_{i,j}$ = O intervalo de compra para o cliente i para a j -ésimo ocasião de compra;

n = O número de anos para a previsão (três anos);

T^* = O período de tempo corrente;

T_i = O número previsto de compras feitas pelo cliente i até o final dos três anos depois de T^* .

Diante da equação, é possível constatar que a alocação dos recursos de *marketing* é assumida como sendo feita anualmente, e essa alocação ocorre no início do ano. Assim, a alocação de custos no primeiro ano não precisa ser descontada, a alocação de custos do segundo ano precisa ser descontada apenas um ano e assim por diante. Isso justifica o denominador do custo de *marketing* com o expoente $t - 1$.

As variáveis estimadas que determinam a previsão do *CLV* são $\hat{Q}_{i,j}$ e $\hat{t}_{i,j}$, que são determinadas por parâmetros específicos ($\beta_{i,p}$) de cada cliente i referentes às p variáveis de decisão de *marketing* (ex.: frequência de contatos por meio de venda pessoal e frequência de contatos por meio de telefone e *email*) e específicas dos clientes (ex.: *cross-buying*, comunicação bidirecional, retornos dos clientes e frequência de contatos pela *WEB*). Para maiores detalhes a respeito da modelagem utilizada, ver Venkatesan, Kumar e Bohling (2007).

O lado esquerdo do sinal de subtração representa as entradas de caixa providas pelo cliente i , enquanto o lado direito da subtração representa as saídas de caixa causadas pelos contatos de *marketing* da empresa com o cliente i . Com isso, portanto, tem-se um modelo de *CLV* no modo *bottom-up*, ou seja, é possível, assim, prever o valor futuro de cada cliente individualmente, sem que se tenha um valor médio igual para todos os clientes. Essa forma de mensurar o *CLV* traz um enorme avanço para o gerenciamento das atividades de *marketing*, que podem ser planejadas com base na alocação de recursos por cliente ou por segmentos de clientes.

O *CE*, por sua vez, não foi abordado no presente modelo, o que implica que seja possível calcular apenas o valor da base de clientes atuais de uma empresa a partir do somatório dos *CLVs* desses clientes. Embora o *CE* não seja o foco de análise desses autores, com o modelo por eles desenvolvido, podemos, portanto, ter no máximo um modelo estático

de *CE*, pois não foi apresentada nenhuma forma de prever os retornos referentes a possíveis aquisições de novos clientes.

Além disso, o modelo trata de forma determinística da obtenção dos valores de variáveis como a margem de contribuição trazida pelo cliente *i* e o número de contatos de *marketing* com o cliente *i*. Isso faz com que o modelo, embora já seja bastante sofisticado, ainda deixe de considerar a natureza estocástica dessas duas variáveis.

2.3.11 Kumar *et al.* (2008)

O modelo proposto por Kumar *et al.* (2008), conforme já citado, possui uma lógica de cálculo semelhante ao modelo de Venkatesan, Kumar e Bohling (2007), no entanto esse modelo diferencia-se justamente pela proposta de estimar também a margem de contribuição trazida pelo cliente *i* e o número de contatos de *marketing* com o cliente *i*. Além disso, o modelo é simplificado à medida que, em vez de estimar a quantidade de compra e o intervalo de compra de cada cliente, se propõe a previsão de uma probabilidade de compra futura do cliente *i*, que, multiplicada pela margem de contribuição total de cada período trazida pelo cliente *i*, permite que seja acessada a entrada de caixa futura gerada pelo cliente *i*. Ao considerar a probabilidade futura de compra, o modelo leva em conta também a probabilidade de que o cliente *i* não compre da empresa focal. A estimação dessas variáveis também é realizada por meio de estatística bayesiana.

O modelo de *CLV* proposto é

$$CLV_i = \sum_{j=T+1}^{T+36} \frac{p(\text{Buy}_{ij} = 1) * \widehat{CM}_{ij}}{(1+r)^{j-T}} - \frac{\widehat{MT}_{ij} * \overline{MC}}{(1+r)^{j-T}} \quad \text{Equação 10}$$

Onde:

CLV_i = Valor vitalício do cliente *i*;

$p(\text{Buy}_{ij})$ = Probabilidade de compra prevista de que o cliente *i* comprará no período de tempo *j*;

CM_{ij} = Margem de contribuição prevista para o cliente *i* no período de tempo *j*;

MT_{ij} = Nível de contatos de *marketing* previsto direcionado ao cliente *i* no período de tempo *j*;

\overline{MC} = Custo médio para um único contato de *marketing*;

j = índice para os períodos de tempo, em meses no caso do estudo realizado pelos autores;

T = Demarca o final do tempo total de calibração ou observação;

r = Taxa de desconto, mensal no caso do estudo realizado pelos autores;

Os autores determinam, então, que as variáveis que devem ser estimadas para a mensuração do *CLV* do cliente i são: $p(\text{Buy}_{i,j} = 1)$, CM_{ij} e MT_{ij} . Essas variáveis são preditas por meio das seguintes regressões que são correlacionadas entre si:

$$\log(MT_{ij}) = \alpha_{1i} + \mathbf{x}_{1ij}^T \boldsymbol{\beta}_1 + \mathbf{u}_{1ij} \quad \text{Equação 11}$$

$$\text{Buy}_{ij}^* = \alpha_{2i} + \mathbf{x}_{2ij}^T \boldsymbol{\beta}_2 + \mathbf{u}_{2ij} \quad \text{Equação 12}$$

$$CM_{ij}^* = \alpha_{3i} + \mathbf{x}_{3ij}^T \boldsymbol{\beta}_3 + \mathbf{u}_{3ij} \quad \text{Equação 13}$$

Essas três variáveis são determinadas pelos interceptos (α_{1i} , α_{2i} e α_{3i}), que são distintos entre cada uma das três regressões e entre cada um dos clientes considerados. O intercepto é provido pelas variáveis preditoras (x_{1ij} , x_{2ij} e x_{3ij}), que são divididas em características de troca (definem e descrevem as características transacionais entre o cliente e a firma) e características firmográficas (que captam a heterogeneidade entre cada cliente). Exemplos desses tipos de características serão apresentados a seguir, uma vez que essas características são também utilizadas por Kumar e Shah (2009). Esse modelo, porém, não controla a heterogeneidade entre clientes por meio dos coeficientes angulares das regressões (β_1 , β_2 e β_3) que variam somente entre as regressões, sendo os mesmos entre todos os clientes.

Sendo assim, o modelo proposto peca neste quesito, porém é importante ressaltar sua importância em conseguir estimar o *CLV* no modo *bottom-up*, considerando a natureza estocástica das principais variáveis que impactam na estimação do *CLV* e ainda fazendo tal mensuração por meio de métodos bayesianos. O *CE*, porém, novamente não é o foco desses autores, o que faz com que esse modelo também não apresente formas para mensurar o valor de clientes futuros que serão adquiridos pela firma.

2.3.12 Kumar e Shah (2009)

Dada a importância dos modelos de Kumar *et al.* (2008) e de Kumar e Shah (2009) para o presente trabalho e considerando que o segundo é uma sequência do primeiro, o modelo mais recente é decomposto em detalhes a seguir e, após, as considerações pertinentes acerca do mesmo são tecidas. O *CLV* é calculado a partir da seguinte equação:

$$CLV_i = \sum_{j=T+1}^{T+36} \frac{p(\text{Buy}_{ij} = 1) * \widehat{CM}_{ij}}{(1+r)^{j-T}} - \frac{\widehat{MT}_{ij} * \overline{MC}}{(1+r)^{j-T}} \quad \text{Equação 14}$$

Onde:

CLV_i = Valor vitalício do cliente i ;

$p(\text{Buy}_{ij} = 1)$ = Probabilidade de compra prevista de que o cliente i comprará no período de tempo j ;

CM_{ij} = Margem de contribuição prevista para o cliente i no período de tempo j ;

MT_{ij} = Nível de contatos de *marketing* previsto direcionado ao cliente i no período de tempo j ;

\overline{MC} = Custo médio para um único contato de *marketing*;

j = índice para os períodos de tempo, em meses no caso do estudo realizado pelos autores;

T = Demarcação do final do tempo total de calibração ou observação;

r = Taxa de desconto, mensal no caso do estudo realizado pelos autores.

Dada a equação principal para o cálculo do *CLV*, autores determinam, então, que as variáveis que devem ser estimadas para a mensuração do *CLV* do cliente i são: $p(\text{Buy}_{i,j} = 1)$, CM_{ij} e MT_{ij} . Essas variáveis são preditas por meio das seguintes regressões, que são correlacionadas entre si:

$$\log(1 + MT_{ij}) = \alpha_{1i} + x_{1ij}^T \beta_{1i} + u_{1ij} \quad \text{Equação 15}$$

$$\text{Buy}_{ij}^* = \alpha_{2i} + x_{2ij}^T \beta_{2i} + u_{2ij} \quad \text{Equação 16}$$

$$CM_{ij}^* = \alpha_{3i} + x_{3ij}^T \beta_{3i} + u_{3ij} \quad \text{Equação 17}$$

Nessas regressões, x_{1ij} , x_{2ij} e x_{3ij} são os vetores das variáveis preditoras; α_{1i} , α_{2i} e α_{3i} são os interceptos específicos para cada cliente; β_{1i} , β_{2i} e β_{3i} são os vetores correspondentes aos coeficientes específicos para cada cliente e para cada variável preditora; e u_{1ij} , u_{2ij} e u_{3ij} são os termos de erro. As variáveis preditoras são as mesmas para as equações (15, 16 e 17) e são obtidas a partir do banco de dados existente sobre as variáveis de troca entre o cliente e a empresa. Segue a lista com todas as variáveis inicialmente consideradas para o cálculo do modelo:

- (a) Número de contatos de *marketing* defasado um período;
- (b) Quantidade de compras em um período;
- (c) Intervalo de tempo em meses desde a última compra;
- (d) Margem de contribuição defasada um período;
- (e) *Share-of-wallet*;
- (f) Número de canais utilizado pelo cliente para transacionar em um dado ano;
- (g) Quantidade de categorias de produtos diferentes compradas pelo cliente em um dado ano;
- (h) Valor monetário dos produtos devolvidos pelo cliente em um dado ano;
- (i) Valor monetário dos créditos recebidos pelo cliente devido a recomendações dadas em um dado ano.

Dado que a variável incidência de compra é uma variável binária, Buy_{ij} (0 = não comprou; 1 = comprou), para que seja possível obter uma probabilidade de compra representada por uma variável contínua entre 0 e 1, os autores fazem uso do Buy_{ij}^* , que representa a utilidade latente de compra. Quando Buy_{ij}^* é maior do que zero, o cliente compra da firma. Como a variável latente não é observada, ela é mapeada com base na variável Buy_{ij} da seguinte forma:

$$Buy_{ij}^* > 0, se Buy_{ij} = 1$$

$$Buy_{ij}^* \leq 0, se Buy_{ij} = 0$$

A variável CM_{ij} , por sua vez, apresenta *missing values* no banco de dados, pois, quando $Buy_{ij} = 0$, não há compra e, conseqüentemente, não há margem de contribuição observada. Isso não significa, porém, que o cliente necessariamente não comprou de ninguém. Sendo assim, os autores consideram que, quando o cliente não comprou da firma focal, ele

pode ter comprado de um concorrente. Diante disso, para controlar o valor que pode ter sido perdido pela firma focal nas ocasiões em que o cliente possa ter comprado dos concorrentes, os autores imputam valores para os *missing values* da variável margem de contribuição. Esses valores são obtidos pelo produto da margem de contribuição média de um cliente específico pela parcela referente aos concorrentes do *share-of-wallet* desse mesmo cliente. A nova variável é chamada de CM_{ij}^* .

A função de máxima verossimilhança correspondente a esse modelo é apresentada a seguir:

Equação 18

$L(\mathbf{MT}, \mathbf{Buy}, \mathbf{CM}) \propto$

$$\prod_{i=1}^N \prod_{j=1}^T \Pr(\mathbf{Buy}_{ij}^* \leq \mathbf{0}, \mathbf{MT}_{ij})^{1-\mathbf{Buy}_{ij}} * \Pr(\mathbf{CM}_{ij}^* = \mathbf{CM}_{ij}, \mathbf{Buy}_{ij}^* > \mathbf{0}, \mathbf{MT}_{ij})^{\mathbf{Buy}_{ij}}$$

As equações 15, 16 e 17 são relacionadas a um mesmo cliente e, assim, são naturalmente correlacionadas. A formação das equações 15, 16 e 17 é similar à estrutura de um modelo de *SUR* (*Seemingly Unrelated Regressions*), de forma que as variáveis preditoras nessas equações precisam ser as mesmas. Diante disso, os autores assumem que a estrutura de covariância dos erros dessas três equações é como segue:

$$\begin{bmatrix} u_{1ij} \\ u_{2ij} \\ u_{3ij} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \right) = N_3(0, \Sigma_y).$$

Esse tipo de estrutura de covariância permite correlação entre os resíduos. σ_{11} está fixado em 1 para assegurar a identificação do modelo. Além disso, essa estrutura de covariância trata de qualquer dependência não observada entre MT_{ij} , Buy_{ij} e CM_{ij} .

Os interceptos específicos para cada cliente para as equações 15, 16 e 17 são obtidos a partir de uma distribuição normal multivariada:

$$\alpha_i \sim \mathbf{MVN}(\Delta \mathbf{Z}_i, \Sigma_\alpha), \text{ onde}$$

\mathbf{Z}_i = um vetor $p \times 1$ de características do cliente,

Δ = uma matriz $3 \times p$ de coeficientes para as características dos clientes,

Σ_α = uma matriz 3 x 3 de variância-covariância,

p = número de características do cliente usadas para captar a heterogeneidade.

Os coeficientes específicos para cada cliente para as equações 15, 16 e 17 são obtidos a partir de uma distribuição normal multivariada:

$$\beta_i \sim \text{MVN}(\psi Z_i, \Sigma_\beta), \text{ onde}$$

Z_i = um vetor p x 1 de características do cliente,

ψ = uma matriz 3 x p de coeficientes para as características dos clientes,

Σ_β = uma matriz 3 x 3 de variância-covariância,

p = número de características do cliente usadas para captar a heterogeneidade.

A variabilidade dos parâmetros é induzida pela heterogeneidade entre os clientes, capturada por meio das variáveis demográficas dos clientes, vetor Z_i , que é constituído por uma quantidade p de características demográficas de cada cliente. Kumar e Shah (2009) utilizaram apenas 3 características (p=3), sem destacar, porém, o porquê dessa escolha. No entanto esses autores apresentam uma lista com todas as características que poderiam ser utilizadas:

- (a) Idade;
- (b) Gênero;
- (c) Estado civil;
- (d) Tipo de domicílio;
- (e) Renda familiar;
- (f) Distância entre a casa e a loja;
- (g) Número de canais utilizados para comprar;
- (h) Informações sobre o cartão de fidelidade.

A partir dessas informações, é possível identificar que os parâmetros α_i e β_i são obtidos por meio dos hiperparâmetros $\Delta, \psi, \Sigma_\alpha$ e Σ_β , que representam as médias e as variâncias de cada distribuição normal multivariada dos parâmetros α_i e β_i . Diante disso, é preciso estabelecer as distribuições *a priori* para cada um desses hiperparâmetros:

Dado que d_Δ representa a dimensão do vetor Δ , então a distribuição *a priori* para Δ é também uma normal multivariada como segue: $\Delta \sim \text{MVN}(\mu_\Delta, \Sigma_\Delta)$, onde μ_Δ é um vetor coluna

de zeros com d_Δ dimensões e $\Sigma_\Delta = I_{d_\Delta}$, sendo que I_{d_Δ} é uma matriz identidade com $d_\Delta \times d_\Delta$ dimensões.

Dado que d_ψ representa a dimensão do vetor ψ , então a distribuição *a priori* para ψ é também uma normal multivariada como segue: $\psi \sim \text{MVN}(\mu_\psi, \Sigma_\psi)$, onde μ_ψ é um vetor coluna de zeros com d_ψ dimensões e $\Sigma_\psi = 100I_{d_\psi}$, sendo que I_{d_ψ} é uma matriz identidade com $d_\psi \times d_\psi$ dimensões.

A distribuição *a priori* para o hiperparâmetro Σ_α é uma Wishart Inversa, de forma que a sua especificação é dada por: $\Sigma_\alpha = \text{IW}(\rho I_{d_\Delta}, \rho)$, onde $\rho = 15$ e I_{d_Δ} é uma matriz identidade com $d_\Delta \times d_\Delta$ dimensões.

A distribuição *a priori* para o hiperparâmetro Σ_β é uma Wishart Inversa, de forma que a sua especificação é dada por: $\Sigma_\beta = \text{IW}(\rho I_{d_\psi}, \rho)$, onde $\rho = 15$ e I_{d_ψ} é uma matriz identidade com $d_\psi \times d_\psi$ dimensões.

Por fim, a distribuição *a priori* para o a variância da distribuição normal dos erros que segue também uma distribuição Wishart Inversa com seguinte especificação: $\Sigma_y = \text{IW}(15I_3, NT)$, onde I_3 é uma matriz identidade com 3×3 dimensões.

A partir dessa estrutura, as distribuições posteriores dos parâmetros são obtidas, possibilitando também estimar a probabilidade de compra, a margem de contribuição e os contatos de *marketing*. Com isso, esses valores referentes a um único cliente são, então, utilizados para calcular o *CLV* (Equação 14).

O *CE*, por sua vez, será calculado a partir da mensuração de duas parcelas que serão posteriormente somadas para obter o valor do *CE* dinâmico. A primeira parte diz respeito ao *CE* dos clientes retidos na base (CE_R), que é calculado a partir do somatório dos *CLVs* obtidos para cada um dos clientes:

$$CE_R = \sum_{i=1}^N CLV_i \quad \text{Equação 19}$$

A outra parcela para o cálculo do *CE* diz respeito à previsão do valor dos clientes que a firma espera adquirir no futuro (CE_A). Para o cálculo do CE_A , os autores partem de um percentual de crescimento anual da base de clientes de acordo com o plano de crescimento da firma (ex.: 3%). Diante disso, chega-se ao número M de clientes que serão adquiridos nos próximos três anos (36 meses). A partir disso, a seguinte equação é aplicada:

$$CE_A = \sum_{k=1}^N \sum_{j=T+1}^{T+36} \frac{CLV_{kj} - A_{kj}}{(1+r)^{j-T}} \quad \text{Equação 20}$$

Onde

CLV_{kj} = O valor vitalício para o cliente k que se espera que seja adquirido no tempo j ;

A_{kj} = O custo de aquisição correspondente ao cliente k que se espera que seja adquirido no tempo j ;

j = índice para os períodos de tempo, em meses, no caso do estudo realizado pelos autores;

T = Demarcação do final do tempo total de calibração ou observação;

r = Taxa de desconto, mensal no caso do estudo realizado pelos autores;

O CLV_{kj} é igual à média dos $CLVs$ dos clientes retidos na base da firma. A_{kj} , por sua vez é igual ao custo de aquisição médio considerado pela firma.

Diante disso, o valor total do CE dinâmico é calculado a partir da seguinte equação:

$$CE = CE_R + CE_A \quad \text{Equação 21}$$

Dada essa explanação do modelo de Kumar e Shah (2009), é possível constatar, primeiramente, que este modelo está baseado quase integralmente no modelo de Kumar *et al.* (2008), porém algumas alterações fazem com que o modelo de 2009 seja mais sofisticado e eficiente do que o modelo de 2008 para estimar CLV :

- (a) Todos os parâmetros (α_{1i} , α_{2i} , α_{3i} , β_{1i} , β_{2i} e β_{3i}) são específicos para cada cliente, captando melhor a heterogeneidade entre cada cliente;
- (b) No modelo de Kumar *et al.* (2008), nos períodos em que os clientes não compravam da empresa focal, como controle, os autores imputaram margens de contribuição nos *missings values* dessa variável a partir de valores aleatórios obtidos por meio de uma distribuição normal gerada a partir dessa variável. No modelo de Kumar e Shah (2009), essa imputação de resultados continuou sendo realizada, porém os valores imputados foram ponderados pelo *share-of-wallet (SOW)* de cada cliente.
- (c) O CE é mensurado, e seu cálculo é o resultado da soma entre o CE dos clientes retidos na base e o CE das futuras aquisições de clientes por parte da firma. Dessa

forma, além de ser um modelo de *CLV* no modo *bottom-up*, o *CE* obtido é dinâmico.

Mesmo diante dessas melhorias, é importante citar Box (1976), que afirma que um bom cientista deve ter flexibilidade e coragem para buscar, reconhecer e explorar os erros dos modelos desenvolvidos, especialmente dos seus próprios modelos, ou seja, o cientista não deve se apaixonar pelo seu próprio modelo. Dessa forma, embora se acredite que o modelo de Kumar e Shah (2009) seja útil, algumas críticas devem ser tecidas:

- (a) Enquanto a mensuração do *CE* dos clientes retidos na base ocorre por meio de um sofisticado modelo de estimação bayesiano do valor vitalício de cada cliente atual, a mensuração do *CE* dos clientes que futuramente serão adquiridos pela empresa é determinístico, baseado no percentual de aquisição de clientes esperado pelo plano de crescimento da firma e no custo de aquisição médio por cliente, que também é aquele considerado pela empresa em seu planejamento.
- (b) As variáveis demográficas consideradas para o cálculo do modelo representam um avanço importante frente aos modelos anteriores de *CLV*, pois melhoram a capacidade de acessar a heterogeneidade entre os clientes. Entretanto, durante os 36 meses utilizados para estimar o *CLV*, essas características demográficas de cada cliente são mantidas constantes, ou seja, supõe-se que não haja variabilidade da renda, do estado civil, do tipo de domicílio, entre outras variáveis, durante o período de três anos. Essa suposição é bastante questionável. Sendo assim, seria recomendável considerar também a variabilidade dessas variáveis ao longo do período de tempo considerado. Dessa forma, seria possível prever, com maior precisão, a flutuação de um mesmo cliente entre os diferentes segmentos considerados.
- (c) Kumar e Shah (2009) citam que as variáveis utilizadas para estimação do *CLV* foram definidas por pesquisa prévia, o que confere maior segurança quanto à relevância dessas variáveis como capazes de serem utilizadas para estimar o *CLV*. Entretanto seria pouco prudente aceitar que essas variáveis são realmente as melhores para estimar o *CLV* ou ainda aceitar que outras variáveis não possam ser utilizadas, principalmente variáveis especificamente relacionadas ao negócio da empresa analisada.

Embora críticas possam e devam ser tecidas, é importante reconhecer que “todos os modelos estão errados, a questão prática é o quão errado ele tem que ser para que não seja

útil” (BOX e DRAPER, 2007). Ou seja, é natural que limitações existam em qualquer modelo de estimação. A questão é saber se o modelo pode ser considerado útil mesmo diante dessas limitações. Sendo assim, o modelo de Kumar e Shah (2009) bem como o modelo de Kumar *et al.* (2008) foram os modelos estudados profundamente no presente trabalho e refletem muito do que foi definido no modelo aplicado, pois, embora limitações existam, suas vantagens, algumas das quais estão descritas a seguir, são capazes de torná-lo útil para gerar uma boa aproximação do *CLV* no modo *bottom-up* e, assim, colocam esse modelo em posição vantajosa em relação aos outros modelos anteriormente apresentados:

- (a) O modelo hierárquico utilizado faz com que a previsão tenha componentes adicionais de variabilidade e seja assim mais confiável (GELMAN *et al.*, 2004). Essa variabilidade adicional é proporcionada pelos hiperparâmetros ψ e Δ e pelas variáveis demográficas utilizados. A utilização desses hiperparâmetros permite o cálculo de interceptos e coeficientes individuais por cliente, o que faz com que, a partir desse modelo, ao contrário de outros modelos, seja possível não só calcular o *CLV* no modo *bottom-up*, como também tornar mais eficiente a capacidade de o modelo diferenciar os clientes entre si.
- (b) Além disso, uma série de variáveis previamente definidas a respeito da relação específica de troca entre cada cliente e a firma forma o conjunto de variáveis preditoras do modelo, o que também proporciona maior capacidade de prever adequadamente as variáveis independentes determinantes do *CLV* (probabilidade de compra, margem de contribuição e contatos de *marketing*).
- (c) Considera a natureza estocástica dos determinantes do *CLV* (probabilidade de compra futura, margem de contribuição e número de contatos de *marketing* com o cliente). Muitos modelos acessam de forma determinística os valores das variáveis que compõem o cálculo do *CLV*, outros, mais sofisticados, entre eles o de Kumar e Shah (2009), buscam obtê-los a partir dos valores que possuem maior probabilidade de ocorrência.
- (d) O *CLV* e suas variáveis independentes são estimadas por meio de métodos bayesianos, o qual combina distribuições *a priori* para os estados de incerteza com dados existentes para obter distribuições posteriores capazes de reduzir as incertezas a respeito dos estados de incerteza e reduzir a possibilidade de erros quando previsões de longo prazo são realizadas (VENKATESAN, KUMAR e BOHLING, 2007);

- (e) O modelo em questão utiliza o modo *bottom-up* para a estimação do *CLV* de cada cliente. Essa característica somada com a sofisticação das estimações realizadas proporciona significativa segurança para um gerente de *marketing* segmentar sua base a partir do valor vitalício de cada cliente e mensurar a produtividade das ações de *marketing*.
- (f) Com a segmentação dos clientes a partir do *CLV*, é possível alocar os recursos de *marketing* com maior precisão, tendo como o objetivo maximizar o valor dos clientes com *CLVs* altos ou médios, eliminar os custos sem retorno, empregados com clientes com *CLVs* baixos ou negativos, e ainda possibilita que a empresa defina o perfil do cliente que em geral possui alto *CLV* e perfil do cliente que em geral possui *CLV* baixo ou negativo.
- (g) Além de gerenciar o valor de cada um dos clientes atuais individualmente, o modelo permite, ainda, o gerenciamento da base de clientes atuais e futuros de forma agregada, ou seja, o gerenciamento do *CE* dinâmico.

3 MÉTODO

Nesta seção, o método utilizado para estimar os *CLVs* de cada cliente e o *CE* é descrito em detalhes. Para tanto, na subseção 3.1, a descrição da empresa e a forma de tratamento dos dados para obtenção da base de dados final são apresentadas. Na subseção 3.2, o modelo construído e aplicado é apresentado. Nas subseções 3.3 e 3.4, a forma de estimação dos parâmetros do modelo e as análises de diagnóstico realizadas são evidenciadas. Na subseção 3.5, a forma de cálculo dos *CLVs* de cada cliente e do *CE* é definida. Na subseção 3.6, é apresentada a maneira pela qual segmentou-se a base de clientes a partir do *CLV*. Por fim, na subseção 3.7, os testes realizados para verificação da capacidade de previsão do modelo são apresentados.

3.1 DADOS UTILIZADOS

3.1.1 Descrição da empresa

A base de clientes utilizada para a realização do presente estudo é proveniente de uma empresa de serviços financeiros em atuação em todo o território brasileiro, a qual aceitou

colaborar na tarefa de executar a construção e aplicação do modelo de *CLV*. O modelo em questão utiliza o modo *bottom-up* para mensurar o *CLV* e, por isso, é preciso dispor de uma base de dados confiável e que contenha informações sobre as variáveis envolvidas no modelo. O modelo aplicado, assim como os modelos de Kumar *et al.* (2008) e de Kumar e Shah (2009), utiliza uma abordagem “*always-a-share*” sobre a relação entre cliente e empresa, o que é mais apropriado para o contexto de relação não contratual. Esse é o caso da empresa de serviços financeiros analisada nesta dissertação, que, embora mantenha contrato formal com seus clientes, estes são reconhecidamente livres para transacionar com outros concorrentes, e o fazem, além de não necessariamente apresentarem uma frequência mensal de compras (frequência média observada entre todos os clientes foi de 1 compra a cada 5 meses).

3.1.2 Base de dados

A base de dados para o total de todos os clientes da empresa foi obtida a partir de cerca de 10.000 arquivos em formatos texto (*.txt), valores separados por vírgula (*.csv) ou excel (*.xls), referentes às informações demográficas e transacionais dos clientes entre os períodos de janeiro de 2011 a maio de 2013. Reunidos todos os arquivos, a base de dados possui aproximadamente 140.000.000 de observações. Para viabilizar o tratamento de tamanha quantidade de dados com confiabilidade, foi necessário utilizar o sistema *MySQL*, que proporciona o tratamento eficiente de grandes bases de dados por meio da linguagem *SQL* (*Structured Query Language*). Mais especificamente, utilizou-se o aplicativo *Sequel Pro*, que é um *software* que proporciona o gerenciamento rápido e fácil de bases de dados *MySQL*. O *MySQL* é um dos diversos Sistemas Gerenciadores de Banco de Dados (SGBD) existentes que utilizam a linguagem *SQL* para realizar os comandos para manipulação da base de dados. O SGBD *MySQL* é reconhecido e utilizado mundialmente por desenvolvedores de *softwares* (ex.: *softwares* ERP e portais) para obter eficiência nas pesquisas e manipulações de grandes bases de dados.

Conforme mencionado, existem ainda outros sistemas de gerenciamento de base de dados que também utilizam a linguagem *SQL*, como o *Microsoft SQL Server*, o *Oracle* e o *PostgreSQL*. No entanto, apesar de existirem discussões de *experts* sobre diferenças de *performance* entre cada uma dessas ferramentas, como todos são sistemas potentes para tratamento de grandes bases de dados, a base de dados utilizada no presente estudo, embora grande o suficiente para não poder ser tratada sem o auxílio desses sistemas, acaba não

levando nenhum desses *softwares* ao limite, fazendo com que as diferenças de *performance* entre um e outro não sejam observadas, nem sejam relevantes para a realização do trabalho. Dessa forma, escolheu-se utilizar o *MySQL* gerenciado a partir da interface do aplicativo *Sequel Pro*.

Diante disso, após um minucioso processo de importação de todos os arquivos para uma única grande base de dados formada por diversas tabelas dentro do *software Sequel Pro*, foi possível manipular os dados, expurgando aqueles inconsistentes e realizando as devidas operações para a definição das variáveis necessárias para a estimação das três variáveis determinantes do *CLV*. Essa etapa do trabalho foi de extrema importância para a obtenção de uma base de dados final que fosse confiável. Esse trabalho foi dividido em quatro etapas:

1. Importação de todos os arquivos para o *software Sequel Pro*;
2. Aprendizado sobre a empresa e sobre a sua base de dados por meio de reuniões lideradas por pessoas da empresa e de dúvidas pontuais sanadas ao longo do processo de tratamento da base de dados. Nesta etapa, foram tomadas decisões, sob orientação dos profissionais da empresa de serviços financeiros, que definiram a forma como as principais variáveis do modelo seriam obtidas. Essas decisões estão descritas a seguir:

- Produtos considerados: A empresa estudada comercializa uma ampla gama de produtos que podem ser agrupados nas seguintes categorias: BMF, Bovespa, Fundos, Renda Fixa, *Initial Public Offering (IPO)*, Clubes de Ações, Seguros e Previdência. Diante da decisão dos profissionais da empresa, as categorias de produtos IPO, Clubes de Ações, Seguros e Previdência não foram incluídas na análise principalmente porque, somadas, as mesmas representam uma parcela muito pequena da receita total da empresa.

Diante disso, definiu-se que as categorias de produtos que comporiam o cálculo da margem de contribuição total de cada cliente seria formada pela soma das margens de contribuição geradas pelas categorias de produtos BMF, Bovespa, Fundos e Renda Fixa. Essas categorias de produtos possuem as seguintes características:

- BMF: Trata-se da negociação principalmente de mercadorias e futuros na bolsa de valores por parte do cliente;
- Bovespa: Trata-se da negociação principalmente de ações e opções na bolsa de valores por parte do cliente;

- Fundos: Trata-se do investimento em fundos de investimento por parte do cliente;
- Renda Fixa: Trata-se do investimento realizado diretamente em Títulos Públicos e Privados de Renda Fixa por parte do cliente.
- Segmentos de clientes considerados:
 - Pessoa Física e Pessoa Jurídica: Os clientes da empresa se dividem entre clientes Pessoa Física e clientes Pessoa Jurídica. Diante da grande disparidade entre as margens de contribuição gerada pelos clientes desses dois segmentos (clientes Pessoa Jurídica possuem movimentação mais elevada), optou-se por simplificar o problema abordado, visando a aplicar o modelo inicialmente apenas com os clientes Pessoa Física. Os clientes Pessoa Física representam 98% do total de clientes e geram 88% da receita total gerada pelos clientes presentes na base de dados final.
 - B2B e B2C: Os clientes da empresa também podem ser divididos em clientes B2B, que são aqueles clientes atendidos por escritórios afiliados, e clientes B2C, que são aqueles clientes atendidos pelos escritórios gerenciados pela matriz. Na análise realizada, os dois tipos de clientes serão considerados.
- Cálculo da margem de contribuição: A definição da forma de cálculo da margem de contribuição foi inteiramente determinada pelas informações providas pelos profissionais da empresa. Duas opções eram possíveis: utilizar um valor fixo de comissão paga aos agentes que atendem os clientes para cada situação dependendo do tipo de cliente e da categoria de produtos ou utilizar o valor aparentemente real de comissões pagas aos agentes. A segunda opção foi inicialmente escolhida e todas as margens de contribuição foram calculadas, entretanto logo descobriu-se que havia informações inconsistentes nos arquivos referentes especificamente às comissões pagas. Diante disso, a primeira opção foi, então, adotada, e as margens de contribuição foram calculadas conforme descrito na Tabela 2:

Tabela 2 – Cálculo da margem de contribuição por categoria de produto

Tipo de Cliente	Tipo de Fundo de Investimento	Fórmula para o cálculo da margem de contribuição
Bovespa		
Cliente B2B	-	Receita bruta (-) Impostos (-) Comissão média do escritório afiliado para a categoria de produtos Bovespa
Cliente B2C	-	Receita bruta (-) Impostos (-) Custos e comissão média da mesa de transações interna
BMF		
Cliente B2B	-	Receita bruta (-) Impostos (-) Comissão média do escritório afiliado para a categoria de produtos BMF
Cliente B2C	-	Receita bruta (-) Impostos (-) Custos e comissão média da mesa de transações interna
Fundos		
Cliente B2B	Fundos Próprios	Receita bruta (-) Impostos (-) Custo médio de gestão do fundo (-) Comissão média do escritório afiliado para a categoria de produtos Fundos
Cliente B2B	Fundos de Terceiros	Receita bruta (-) Impostos (-) Comissão média do escritório afiliado para a categoria de produtos Fundos
Cliente B2C	Fundos Próprios	Receita bruta (-) Impostos (-) Custo médio de gestão do fundo (-) Custos e comissão média da mesa de transações interna
Cliente B2C	Fundos de Terceiros	Receita bruta (-) Impostos (-) Custos e comissão média da mesa de transações interna
Renda Fixa		
Cliente B2B	-	Receita bruta (-) Impostos (-) Comissão média do escritório afiliado para a categoria de produtos Renda Fixa
Cliente B2C	-	Receita bruta (-) Impostos (-) Custos e comissão média da mesa de transações interna

- Apropriação dos custos de *marketing*: Diante da ausência de um histórico da quantidade de contatos de *marketing* realizados com cada cliente, o que seria um importante parâmetro para apropriar os custos de *marketing* com maior precisão, adotou-se a política de relacionamento com o cliente utilizada pela empresa como base para realizar a apropriação dos custos de *marketing* para cada cliente em cada mês do período em análise.

Os custos de *marketing* computados mensalmente pela empresa são os seguintes:

- Custos indiretos: Custos com ações de *marketing* e materiais institucionais repassados aos escritórios afiliados (clientes B2B);
- Custos da matriz: Custos com ações de *marketing* e materiais institucionais realizadas pela matriz (clientes B2C);
- Custos de telefone: Custos dos atendimentos por telefone realizados pela matriz (clientes B2C).

A política de relacionamento com os clientes utilizada pela empresa é a seguinte:

- Clientes inativos (possuem posição financeira (*net*) = 0) não são contatados;
- Clientes ativos com posição financeira (*net*) menor do que “X” reais também não são contatados, esse cliente deve realizar as transações por conta própria;
- Clientes ativos com posição financeira (*net*) maior ou igual a “X” reais e menor do que 6 vezes X reais são contatados bimestralmente;
- Clientes ativos com posição financeira (*net*) maior ou igual a 6 vezes “X” reais são contatados mensalmente.

Diante dessa política, as seguintes ações foram tomadas para alocar os custos de *marketing*:

- Como a empresa não possui capacidade de gerenciar os contatos que os escritórios afiliados realizam para com os clientes, todos os custos indiretos mensais foram igualmente alocados entre os clientes B2B ativos;
- Como os custos da matriz são custos relacionados a ações de *marketing* não realizadas diretamente em relação a determinados clientes (ex.: ações institucionais), todos os custos da matriz mensais foram alocados igualmente para os clientes B2C ativos;

- Como os contatos de telefone seguem a política de relacionamento com o cliente, adotou-se a seguinte regra para alocar os custos de *marketing* para os clientes B2C ativos: (a) peso 2 foi atribuído aos clientes com posição financeira (*net*) maior ou igual a 6 vezes “X” reais; (b) peso 1 foi atribuído para os clientes com posição financeira (*net*) maior ou igual a “X” reais e menor do que 6 vezes “X” reais e (c) peso 0 foi atribuído aos clientes com posição financeira (*net*) menor do que “X” reais. Diante disso, os custos de telefone mensais foram alocados de acordo com o peso que cada cliente possuía em cada mês.
 - Os clientes inativos em um determinado mês, sejam eles B2B ou B2C, não receberam nenhum custo de *marketing*.
3. Com as principais informações sobre a empresa, o seu negócio e a sua base de dados absorvidas e formalmente descritas, o trabalho bastante denso de tratamento da base de dados pôde ser realizado. Nesse processo, aqueles dados comprovadamente inconsistentes foram expurgados e as informações consistentes que seriam necessárias para o cálculo do modelo foram identificadas, calculadas quando necessário e reunidas em apenas duas tabelas finais:
- a. Na primeira tabela, foram reunidas todas aquelas informações longitudinais com variação mensal ou anual, dependendo da variável (ex.: margem de contribuição, custos de *marketing*, etc.). Esta tabela possui informações sobre 29 meses de transações, de janeiro de 2011 a maio de 2013, para cada um dos 90.387 clientes, gerando uma base de dados com 2.621.223 linhas;
 - b. Na segunda tabela, foram reunidas todas aquelas informações constantes para cada cliente (ex.: salário, sexo, frequência de compra, etc.), totalizando uma base de dados com 90.387 linhas, referentes aos 90.387 clientes.
4. Por fim, um trabalho de análise descritiva dos dados foi realizado com o objetivo de assegurar a consistência das informações obtidas (ex.: conferir se os resultados dos cálculos realizados estavam corretos, conferir a receita anual a partir da base de dados frente à receita publicamente divulgada pela empresa, entre outras análises).

Após todas as operações necessárias realizadas, chegou-se aos valores finais das três variáveis dependentes determinantes do *CLV*, bem como de uma lista de variáveis candidatas

a variáveis preditoras de cada uma dessas três variáveis dependentes. Essas variáveis candidatas a preditoras foram definidas com base nas variáveis utilizadas por Kumar *et al.* (2008) e por Kumar e Shah (2009), bem como no conhecimento adquirido sobre o negócio da empresa nas discussões com os seus profissionais. A relação das variáveis presente nas duas tabelas da base de dados final é apresentada na Tabela 3.

Por fim, frente à grande base de dados, à falta de capacidade computacional disponível e à necessidade ainda pendente de otimização do complexo código desenvolvido para a estimação dos parâmetros e previsão das variáveis determinantes para o cálculo do *CLV*, optou-se por trabalhar com uma amostra de clientes da empresa. Sendo assim, decidiu-se primar pela capacidade de aplicação do modelo construído, que exigiu a realização de uma elevada quantidade de testes. Caso amostras maiores fossem utilizadas, o elevado número de vezes que o modelo precisou ser calculado não seria possível. Diante disso, uma amostra de 100 clientes foi extraída aleatoriamente da base de dados completa e utilizada para a aplicação do modelo. Como existem variáveis com defasagem de até dois períodos consideradas, os dois primeiros períodos de cada cliente (janeiro e fevereiro de 2011) precisaram ser excluídos da análise, pois essas variáveis com defasagem apresentam *missing values* para o primeiro e segundo períodos. Sendo assim, a amostra final utilizada para a aplicação do modelo construído é formada por 100 clientes e 27 meses para cada cliente, totalizando 2700 observações.

3.2 DESCRIÇÃO DO MODELO CONSTRUÍDO E APLICADO

O modelo em questão realiza as estimações dos parâmetros com base em estatística bayesiana. A teoria de decisão bayesiana é ideal para a aplicação em problemas nos quais a decisão precisa ser feita com a modelagem de incerteza. A teoria de decisão bayesiana postula que existem ações (ex.: definir variáveis de decisão de *marketing* – *CLV* –, selecionar clientes, entre outras) que uma firma precisa realizar, as quais possuem estados de incerteza (ex.: probabilidade de compra, margem de contribuição e número de contatos de *marketing*), e que a combinação entre as ações e os estados de incerteza resultam em consequências (ex.: lucros). O Teorema de Bayes combina, então, distribuições *a priori* para os estados de incerteza com dados existentes para obter distribuições posteriores capazes de reduzir as

Tabela 3 – Variáveis consideradas

Variável	Descrição
Código do cliente	Código único para cada cliente.
MC Total*	Margem de contribuição por cliente gerada em cada mês pelas categorias de produtos Bovespa, BMF, Fundos e Renda Fixa.
MC Total <i>Lag</i> 1	Margem de contribuição por cliente gerada em T-1 pelas categorias de produtos Bovespa, BMF, Fundos e Renda Fixa.
MC Total <i>Lag</i> 2	Margem de contribuição por cliente gerada em T-2 pelas categorias de produtos Bovespa, BMF, Fundos e Renda Fixa.
<i>Buy</i> *	Indica se o cliente transacionou em cada mês, sendo 0 = não comprou e 1 = comprou.
<i>Buy Lag</i> 1	Indica se o cliente transacionou em T-1, sendo 0 = não comprou e 1 = comprou.
<i>Buy Lag</i> 2	Indica se o cliente transacionou em T-2, sendo 0 = não comprou e 1 = comprou.
Custos de <i>Marketing</i> *	Custos de <i>marketing</i> apropriado por cliente em cada mês.
Custos de <i>Marketing Lag</i> 1	Custos de <i>marketing</i> apropriado por cliente em T-1.
Custos de <i>Marketing Lag</i> 2	Custos de <i>marketing</i> apropriado por cliente em T-2.
<i>Cross-buying</i>	Quantidade de categorias de produtos diferentes compradas por cliente em cada mês.
<i>Net</i>	Valor que cada cliente possui investido em cada mês. É a posição do cliente.
<i>Net</i> médio anual	Valor médio que cada cliente possui investido em cada ano. É a posição média do cliente por ano.
Total de dias com transação	Quantidade dias de cada mês nos quais cada cliente transacionou apenas os produtos BMF e Bovespa.
Recência desde última compra	Quantidade de meses, calculado para cada cliente e a cada mês, desde a última compra feita pelo cliente.
Períodos com compra	Quantidade de meses em que cada cliente realizou compras durante todo o período considerado.
Idade	Idade de cada cliente.
Sexo	Sexo de cada cliente.
Estado civil	Indica se o cliente é casado ou solteiro.
Salário	Salário declarado pelo cliente em seu cadastro.
Aplicações	Valor total que o cliente possui em aplicações, independentemente do fornecedor, declarado pelo cliente em seu cadastro.

<i>Net médio</i>	Valor médio que o cliente possui investido em todo o período de análise. É a posição média do cliente em todo o período de análise.
<i>Share-of-wallet</i>	Indica o valor do <i>share-of-wallet</i> do cliente, calculado através da razão entre o valor de aplicações com a empresa focal (<i>Net médio</i>) e o valor total de aplicações que o cliente faz com todos os seus fornecedores declarado em seu cadastro.
BMF é categoria majoritária	Indica se a categoria de produtos BMF é a categoria majoritária adquirida pelo cliente.
Bovespa é categoria majoritária	Indica se a categoria de produtos Bovespa é a categoria majoritária adquirida pelo cliente.
Fundos é categoria majoritária	Indica se a categoria de produtos Fundos é a categoria majoritária adquirida pelo cliente.
Renda Fixa é categoria majoritária	Indica se a categoria de produtos Renda Fixa é a categoria majoritária adquirida pelo cliente.
Perfil de risco	Indica o perfil de risco do cliente, sendo 1 = Agressivo (adquire as majoritariamente as categorias de produtos BMF ou Bovespa) e 0 = Conservador (adquire as majoritariamente as categorias de produtos Fundos ou Renda Fixa).
Tempo de relacionamento	Indica se o cliente estava na base de clientes no início do período em análise (Jan/2011) ou não.
Tipo de cliente	Indica se o cliente é B2B (atendido por escritórios afiliados) ou B2C (atendido pela matriz).
Recência de compra	Quantidades de meses desde a última compra que o cliente realizou, tendo como mês base o último mês analisado.
Frequência de compra	Quantidade média de meses entre uma compra e outra, calculada a por meio da razão entre o total de períodos em análise e o total de períodos no qual o cliente realizou compra.

*Variáveis dependentes do modelo estimado

incertezas a respeito dos estados de incerteza e reduzir a possibilidade de erros quando previsões de longo prazo são realizadas (VENKATESAN, KUMAR e BOHLING, 2007). Dessa forma, as ações de *marketing* podem ser tomadas de forma mais precisa e os resultados serão superiores – nesse caso, o impacto sobre o lucro.

Nos últimos anos, tem ocorrido um grande aumento do uso de métodos bayesianos em *marketing*. Para Allenby (2014) é possível verificar a existência de uma revolução bayesiana na pesquisa de *marketing*. Para esse autor, isso ocorre, entre outros argumentos, devido à capacidade de geração de boas estimativas até quando se tem uma quantidade de informações reduzida para cada indivíduo analisado e devido à capacidade de se trabalhar com modelos com alta quantidade de parâmetros. Esses argumentos reforçam a ideia de que os métodos bayesianos são particularmente apropriados para a orientação da decisão em relação aos problemas de *marketing*. Enquanto o conceito promissor dos métodos bayesianos é há muito tempo reconhecido, é a recente popularidade que emana do surgimento de divisores de águas como a disponibilidade de capacidade computacional e os esforços para desenvolver modelos novos que tornou esses métodos atrativos para muitos dos problemas de *marketing*.

Em particular, os modelos bayesianos hierárquicos (como o aplicado no presente estudo), baseados em uma sequência de distribuições condicionais, oferecem enorme flexibilidade e modularidade e são bastante úteis para a área de *marketing*. O *marketing* precisa reconhecer que os clientes são diferentes em suas preferências por produtos e que a firma precisa explicitamente levar isso em consideração para determinar ações ótimas. Diante disso, os métodos bayesianos, especialmente aqueles que envolvem modelos hierárquicos, são importantes ferramentas para tornar a realização dessa tarefa cada vez mais eficiente (ROSSI e ALLENBY, 2003). Entre as justificativas que sustentam essa afirmação, a possibilidade de trabalhar com uma quantidade elevada de parâmetros, chegando-se a parâmetros específicos para cada indivíduo analisado, é uma das mais relevantes.

A inferência bayesiana é conduzida sob regras formais da teoria de probabilidades. Isso significa que os estimadores bayesianos são consistentes, assintoticamente eficientes e admissíveis. Em geral, portanto, os métodos bayesianos proveem uma ótima aproximação diante do nível de incerteza existente ou, em outras palavras, da quantidade de informação provida pelo modelo (ROSSI e ALLENBY, 2003). Sendo assim, com a certeza de estar fazendo uso de uma poderosa técnica de inferência e de apoio à decisão, o modelo que será aplicado está baseado em um modelo bayesiano hierárquico, que envolve a estimação de parâmetros com base em hiperparâmetros com distribuições *a priori* definidas.

O processo de construção do modelo passou por uma etapa de compreensão dos modelos de Kumar *et al.* (2008) e de Kumar e Shah (2009), na qual se buscou compreender cada parte desses modelos e a forma como foram implementados. Durante essa etapa, constatou-se teoricamente a forma de implementação desses modelos, no entanto os dois artigos publicados não trouxeram todas as informações sobre o método utilizado e não foi possível a obtenção dessas informações junto aos autores. As condicionais completas que foram utilizadas para a estimação dos parâmetros, por exemplo, não foram publicados nos artigos. Diante dessa situação, uma vasta pesquisa na literatura sobre modelos bayesianos hierárquicos e sobre modelos bayesianos com *Seemingly Unrelated Regressions (SUR)* foi conduzida. Diante dessa pesquisa, foi possível ter entendimento sobre o tema para assegurar que os modelos Kumar *et al.* (2008) e de Kumar e Shah (2009) tiveram suas raízes no modelo de Cowles, Carlin e Connet (1996) e nas condicionais completas utilizadas por esses autores. Sendo assim, diante da revisão teórica realizada e, principalmente, frente aos modelos de Kumar *et al.* (2008) e Kumar e Shah (2009), bem como o modelo de Cowles, Carlin e Connet (1996), o modelo aplicado no presente estudo pôde ser construído. O *software* R foi utilizado para implementar o modelo.

Seguindo os passos de Kumar e Shah (2009), a equação principal para o cálculo do *CLV* individual de cada cliente é o cálculo do valor presente líquido dos fluxos de caixa de 36 períodos futuros. Para cada cliente, os fluxos de caixa são formados pela margem de contribuição estimada menos os custos de *marketing* estimados que a empresa empregou. Dessa forma, a equação a seguir resume a forma de cálculo do *CLV* para cada cliente.

$$CLV_i = \sum_{j=T+1}^{T+36} \frac{p(\text{Buy}_{ij} = 1) * \widehat{CM}_{ij}}{(1+r)^{j-T}} - \frac{\widehat{MC}_{ij}}{(1+r)^{j-T}} \quad \text{Equação 22}$$

Onde:

CLV_i = Valor vitalício do cliente i ;

$p(\text{Buy}_{ij} = 1)$ = Probabilidade de compra prevista de que o cliente i comprará no período de tempo j ;

CM_{ij} = Margem de contribuição prevista para o cliente i no período de tempo j ;

MC_{ij} = Custo de *marketing* previsto direcionado ao cliente i no período de tempo j ;

j = Índice para os períodos de tempo, em meses no caso do presente estudo;

T = Demarcação do final do tempo total de calibração ou observação;

r = Taxa de desconto, mensal no caso do presente estudo (1,171% ao mês, equivalente a uma taxa de 15% ao ano).

A partir desta equação, pode-se evidenciar que a única diferença em relação àquela utilizada por Kumar e Shah (2009) é simplificação da forma de estimação dos custos de *marketing*. Kumar e Shah (2009) optaram por estimar a quantidade de contatos de *marketing* que a empresa realizava com cada cliente e, então, multiplicar esse valor estimado por um valor médio (fixo) referente ao custo de um único contato de *marketing*. Como a empresa analisada no presente estudo não possui sistema de *CRM* capaz de evidenciar de forma confiável a quantidade de contatos de *marketing* realizados, optou-se por alocar os custos de *marketing* contabilizados diretamente para cada cliente de acordo com a política de contatos de *marketing* utilizada pela empresa. Dessa forma, a variável estimada foi o custo de *marketing* que a empresa empregou alocado para cada cliente.

A partir do que foi utilizado por Kumar e Shah (2009), dada a equação principal para o cálculo do *CLV*, as variáveis que devem ser estimadas para a mensuração do *CLV* do cliente i são: $p(\text{Buy}_{i,j} = 1)$, CM_{ij} e MC_{ij} . Essas variáveis são preditas por meio das seguintes regressões, que são correlacionadas entre si:

$$MC_{ij} = \alpha_{1i} + \mathbf{x}_{1ij}^T \beta_1 + u_{1ij} \quad \text{Equação 23}$$

$$\text{Buy}_{ij}^* = \alpha_{2i} + \mathbf{x}_{2ij}^T \beta_2 + u_{2ij} \quad \text{Equação 24}$$

$$CM_{ij}^* = \alpha_{3i} + \mathbf{x}_{3ij}^T \beta_3 + u_{3ij} \quad \text{Equação 25}$$

Nessas regressões, x_{1ij} , x_{2ij} e x_{3ij} são os vetores das variáveis preditoras; α_{1i} , α_{2i} e α_{3i} são os interceptos específicos para cada cliente; β_1 , β_2 e β_3 são os vetores correspondentes aos coeficientes para cada variável preditora; e u_{1ij} , u_{2ij} e u_{3ij} são os termos de erro. As variáveis preditoras para as equações (23, 24 e 25) são obtidas a partir do banco de dados existente sobre as variáveis de troca entre o cliente e a empresa e as variáveis demográficas de cada cliente. Como será visto a seguir, a estrutura de regressões aparentemente não relacionadas (*Seemingly Unrelated Regressions – SUR*) utilizada permite que diferentes variáveis preditoras sejam utilizadas para cada regressão.

A variável MC_{ij} , por se tratar de uma variável observada e sem *missing values*, é incluída de forma direta no modelo. Diante disso, não há necessidade de geração de variável latente referente a essa variável.

A variável incidência de compra (Buy_{ij}) é uma variável binária (0 = não comprou; 1 = comprou). Para que seja possível obter uma probabilidade de compra representada por uma variável contínua, conforme Kumar *et al.* (2008) e Kumar e Shah (2009), faz-se uso do Buy_{ij}^* , que representa a utilidade latente de compra. Quando Buy_{ij}^* é maior do que zero, o cliente compra da firma. Como a variável latente não é observada, ela é mapeada com base na variável Buy_{ij} da seguinte forma e gerada a partir de sua distribuição condicional completa:

$$Buy_{ij}^* > 0, se Buy_{ij} = 1$$

$$Buy_{ij}^* \leq 0, se Buy_{ij} = 0$$

A variável CM_{ij} , por sua vez, apresenta *missing values* no banco de dados, pois, quando $Buy_{ij} = 0$, não há compra e, conseqüentemente, não há margem de contribuição observada. Isso não significa, porém, que o cliente necessariamente não comprou de ninguém. Sendo assim, é prudente considerar que, quando o cliente não comprou da firma focal, ele pode ter comprado de um concorrente ou não ter comprado de ninguém. Diante disso, para controlar o valor que pode ter sido perdido pela firma focal nas ocasiões em que o cliente possa ter comprado dos concorrentes, optou-se por imputar valores para os *missing values* da variável margem de contribuição. Essa geração de valores para os *missing values* do CM_{ij} é feita da seguinte forma:

1° - Calcula-se o valor da frequência média de compra de cada cliente;

2° - Calcula-se, para cada cliente e a cada período, a recência desde a última compra;

3° - Nos períodos em que houver *missing value* e que a recência de compra for maior do que a frequência média de compra do cliente, então um valor negativo gerado a partir da condicional completa para geração do CM^*_{ij} é imputado;

4° - Nos períodos em que houver *missing value* e que a recência de compra for menor do que a frequência média de compra do cliente, então o valor zero é imputado;

Kumar *et al.* (2008) utilizaram regra semelhante, também imputando valores negativos, pois, dessa forma, considera-se a margem de contribuição perdida para os concorrentes. Além disso, é importante ressaltar que os valores imputados nos *missing values*

da margem de contribuição são considerados apenas para a etapa de geração dos parâmetros do modelo. Na etapa de previsão dos valores futuros, os valores sem os dados imputados são considerados.

A geração da variável latente Buy_{ij}^* e dos *missing values* para obtenção do CM_{ij}^* é uma etapa do modelo denominada *data augmentation* (TANNER e WONG, 1987). As condicionais completas para geração dos valores para essas duas variáveis foram obtidas a partir das condicionais completas aplicadas por Cowles, Carlin e Connet (1996). Adiante, na seção em que as etapas para geração da variável aumentada e dos parâmetros forem evidenciadas, todas as condicionais completas implementadas através do amostrador de Gibbs serão definidas.

A estrutura geral do modelo é derivada da estrutura de regressões aparentemente não relacionadas (*SUR*) de Zellner (1971). “O modelo *SUR*, de certo modo, é uma generalização do tradicional modelo de regressão multivariada, de forma que a matriz X , que aparece em cada equação do modelo tradicional, pode ser diferente no modelo *SUR*” (ZELLNER, 1971, pág. 241). Essa afirmação indica que no modelo *SUR*, diferentemente do modelo tradicional de regressão multivariada, as variáveis independentes de cada equação não precisam ser as mesmas. De acordo com Zellner (1971, pág. 241), a estrutura genérica para a construção do modelo *SUR* é a seguinte:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_m \end{pmatrix} = \begin{pmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_m \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_m \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_m \end{pmatrix}$$

Sendo assim, assume-se que a estrutura de covariância dos erros de cada uma das três regressões pode ser modelada da seguinte forma:

$$\begin{bmatrix} u_{1ij} \\ u_{2ij} \\ u_{3ij} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \right) = N_3(0, \Sigma_y)$$

Esse tipo de estrutura de covariância permite correlação entre os resíduos. σ_{11} está fixado em 1 para assegurar a identificação do modelo. Além disso, essa estrutura de covariância trata de possíveis dependências não observadas entre MC_{ij} , Buy_{ij} e CM_{ij} .

Sendo assim, diante da utilização do amostrador de Gibbs e da geração de dados completos por meio do *data augmentation*, a função de máxima verossimilhança pode ser simplificada para um modelo linear multivariado, conforme segue:

$$L(\alpha, \beta, \Sigma_y, MC, Buy^*, CM^*)$$

$$\propto |\Sigma_y|^{-\frac{Nn}{2}} \exp\left(-\frac{1}{2} \text{tr} \sum_{i=1}^N \sum_{j=1}^n \Sigma_y^{-1} (y^*_{ij} - x_{ij}\beta - \alpha_i) \times (y^*_{ij} - x_{ij}\beta - \alpha_i)^T\right)$$

Equação 26

Onde:

$$y^*_{ij} = \begin{bmatrix} MC_{ij} \\ Buy^*_{ij} \\ CM^*_{ij} \end{bmatrix}, \quad \alpha_i = \begin{bmatrix} \alpha_{1i} \\ \alpha_{2i} \\ \alpha_{3i} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \quad x_{ij} = \begin{bmatrix} x_{1ij}^T & 0 & 0 \\ 0 & x_{2ij}^T & 0 \\ 0 & 0 & x_{3ij}^T \end{bmatrix},$$

N = número de clientes e n = número de períodos.

Em relação às especificações das *prioris* dos parâmetros, temos o seguinte:

- Para β , a priori não informativa de Jeffreys foi especificada;
- Para Σ_y , a priori não informativa de Jeffreys foi especificada;
- Para α_i , assume-se que α_i são independentes e identicamente distribuídos (iid) e possuem a seguinte *priori* informativa:

$$\circ \alpha_i = N(0, \Sigma_\alpha), \quad [\Sigma_\alpha] = IW((\rho R)^{-1}, \rho).$$

W refere-se à distribuição Wishart. O valor dos parâmetros para a distribuição Wishart do hiperparâmetro Σ_α , objetivando que os dados dominem a *priori* na determinação da posteriori de Σ_α , são os seguintes:

$$\circ \rho = N/20$$

$$\circ R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3.3 ESTIMAÇÃO DOS PARÂMETROS

Diante da elevada dimensionalidade do modelo construído, sua solução analítica torna-se inviável. Para resolver tal problema, com a base na verossimilhança evidenciada na equação 26, é possível estimar as distribuições posteriores dos parâmetros do modelo por meio de um dos métodos de *Markov Chain Monte Carlo (MCMC)*, o amostrador de Gibbs (GEMAN e GEMAN, 1984; GELFAND e SMITH, 1990).

O Amostrador de Gibbs é um método de amostragem iterativo de uma Cadeia de Markov, cuja transição de um estado a outro é feita a partir das distribuições condicionais completas (a partir de um vetor de parâmetros θ à posteriori, define-se a condicional completa de um subvetor paramétrico genérico θ como a distribuição deste, dado todos os outros parâmetros ($\theta_{-\theta_i}$) e os dados (Y), que será denotado por $p(\theta_i|\theta_{-\theta_i}, Y)$ (CAMPOS, 2007).

A implementação do amostrador de Gibbs deve seguir os seguintes passos:

$$\text{Passo 1: } \theta_1^{(k)} \sim p_1(\theta_1 | \theta_2^{(k-1)}, \theta_3^{(k-1)}, \dots, \theta_n^{(k-1)}, Y)$$

$$\text{Passo 2: } \theta_2^{(k)} \sim p_2(\theta_2 | \theta_1^{(k)}, \theta_3^{(k-1)}, \dots, \theta_n^{(k-1)}, Y)$$

$$\text{Passo 1: } \theta_3^{(k)} \sim p_3(\theta_3 | \theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_n^{(k-1)}, Y)$$

⋮
⋮
⋮

$$\text{Passo n: } \theta_n^{(k)} \sim p_n(\theta_n | \theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_{n-1}^{(k)}, Y).$$

A realização dos passos 1 até n são suficientes para estimar 1 valor para cada parâmetro θ_i , ou seja, a realização dos passos 1 até n correspondem a 1 iteração. Diante disso, cabe ao pesquisador repetir k vezes os passos 1 até n. Na primeira iteração ($k = 1$), valores iniciais quaisquer devem ser atribuídos para os parâmetros apenas para iniciar as iterações. A partir da segunda iteração, os valores $\theta_1^{(k)}, \theta_2^{(k)}, \theta_3^{(k)} \dots, \theta_n^{(k)}$ da iteração $k - 1$ devem ser os valores iniciais $\theta_1^{(k-1)}, \theta_2^{(k-1)}, \theta_3^{(k-1)}, \dots, \theta_n^{(k-1)}$ da iteração k . Sendo assim, utilizando o amostrador Gibbs, gera-se uma longa realização da cadeia com comprimento (número de iterações) t para cada parâmetro estimado, com $t = b + k * m$, onde b é o número inicial de iterações necessárias para que a cadeia atinja o estado de equilíbrio e que será descartado (*burn-in*), k é o número final das iterações que vão ser usadas na aplicação do método de Monte Carlo e m é o espaçamento entre iterações sucessivas delineado para eliminar a autocorrelação entre as consecutivas iterações (*thining*) (PAULINO, TURKMAN e MURTEIRA, 2003). Segundo esses mesmos autores, não existe uma regra geral para cada um desses parâmetros e cada problema individual tem que ser analisado cuidadosamente para a definição de b , k e m . Ao final, se o valor de k for suficientemente grande para atingir a convergência de cada cadeia de valores gerada para cada parâmetro, de forma que cada cadeia não apresente autocorrelação, então se pode considerar encontradas as distribuições posteriores dos parâmetros, pois as mesmas podem ser consideradas amostras representativas da verdadeira distribuição estacionária subjacente (COWLES e CARLIN, 1996). Adiante,

quando as análises de diagnóstico utilizadas no presente estudo serão apresentadas, os testes utilizados para verificação de convergência das cadeias obtidas serão descritos.

Diante da estrutura construída para o modelo e da utilização do amostrador de Gibbs para a estimação das distribuições posteriores para os parâmetros, as condicionais completas para geração da variável aumentada (CM^*_{ij}) e para cada parâmetro (α_i , Σ_α , β e Σ_y), a partir do trabalho realizado por Cowles, Carlin e Connet (1996), estão definidas a seguir:

Passo 1 – Geração do Buy^*_{ij} :

O passo 1 é realizado fora do amostrador de Gibbs, assumindo-se valores de média igual a 0 e variância igual a 1 para geração do vetor latente utilizado na simulação do amostrador de Gibbs:

$$Buy^*_{ij} = TN_{(-\infty, \infty)}(0, 1)$$

Dado a seguinte regra:

$$Buy^*_{ij} > 0, \text{ se } Buy_{ij} = 1$$

$$Buy^*_{ij} \leq 0, \text{ se } Buy_{ij} = 0$$

Tendo-se o vetor de custos de *marketing* observados (MC_{ij}) e o vetor latente Buy^*_{ij} , que são constantes, a simulação por meio do amostrador de Gibbs é realizada conforme segue:

Passo 2 - Condicional completa para cada *missing value* CM^*_{ij} :

$$\begin{aligned} & \left[CM^*_{ij} \mid \beta_3^{(k-1)}, \alpha_{3i}^{(k-1)}, \sigma_2^{2(k-1)}, \sigma_3^{2(k-1)}, \sigma_{23}^{(k-1)}, Buy^*_{ij} \right] \\ & = TN_{(-\infty, 0)} \left(x_{3ij}^T \beta_3^{(k-1)} + \alpha_{3i}^{(k-1)}, \frac{\sigma_{23}^{(k-1)}}{\sigma_2^{2(k-1)}} \right) \\ & \times \left(Buy^*_{ij} - x_{2ij}^T \beta_2^{(k-1)} + \alpha_{2i}^{(k-1)} \right), \sigma_3^{2(k-1)} - \frac{\sigma_{23}^{(k-1)}}{\sigma_2^{2(k-1)}} \end{aligned}$$

Equação 27

Considerando a regra para imputação de valores para os *missing values* de CM_{ij} evidenciada anteriormente.

Passo 3 – Condicional completa para β :

$$\begin{aligned}
& \left[\beta^{(k)} \mid \Sigma_y^{(k-1)}, \alpha^{(k-1)}, y^{*(k)} \right] \\
&= N \left(\left(\sum_{i=1}^N \sum_{j=1}^n x_{ij}^T \Sigma_y^{-1(k-1)} x_{ij} \right)^{-1} \right. \\
&\quad \times \left. \left(\sum_{i=1}^N \sum_{j=1}^n x_{ij}^T \Sigma_y^{-1(k-1)} \left(y_{ij}^{*(k)} - \alpha_i^{(k-1)} \right) \right) \right), \\
&\quad \left(\sum_{i=1}^N \sum_{j=1}^n x_{ij}^T \Sigma_y^{-1(k-1)} x_{ij} \right)^{-1}
\end{aligned}$$

Equação 28

Passo 4 – Condicional completa para α_i :

$$\begin{aligned}
& \left[\alpha_i^{(k)} \mid \Sigma_y^{(k-1)}, \beta^{(k)}, \Sigma_\alpha^{(k-1)}, \mu_\alpha = 0, y^{*(k)} \right] \\
&= N \left(\left(n \Sigma_y^{-1(k-1)} + \Sigma_\alpha^{-1(k-1)} \right)^{-1} \times \left(\sum_{i=1}^N \Sigma_y^{-1(k-1)} \left(y_{ij}^{*(k)} - x_{ij} \beta^{(k)} \right) \right) \right), \\
&\quad \left(n \Sigma_y^{-1(k-1)} + \Sigma_\alpha^{-1(k-1)} \right)^{-1}
\end{aligned}$$

Equação 29

Passo 5 – Condicional completa para Σ_α :

$$\begin{aligned}
& \left[\Sigma_\alpha^{-1(k)} \mid \alpha^{(k)}, R, \rho, N \right] \\
&= W \left(\left(\sum_{i=1}^N \alpha_i^{(k)} \alpha_i^{(k)T} + (\rho R)^{-1} \right)^{-1}, N + \rho \right) \\
&\Leftrightarrow \left[\Sigma_\alpha^{(k)} \mid \alpha^{(k)}, R, \rho, N \right] \\
&= IW \left(\left(\sum_{i=1}^N \alpha_i^{(k)} \alpha_i^{(k)T} + (\rho R)^{-1} \right)^{-1}, N + \rho \right)
\end{aligned}$$

Equação 30

Passo 6 – Condicional completa para a matriz A (etapa necessária para geração do Σ_y)

$$A^{(k)} = \begin{bmatrix} a_1^{2(k)} & a_{12}^{(k)} & a_{13}^{(k)} \\ a_{12}^{(k)} & a_2^{2(k)} & a_{23}^{(k)} \\ a_{13}^{(k)} & a_{23}^{(k)} & a_3^{2(k)} \end{bmatrix}$$

$$\equiv \sum_{i=1}^N \sum_{j=1}^n (y_{ij}^{*(k)} - x_{ij}\beta^{(k)} - \alpha_i^{(k)}) \times (y_{ij}^{*(k)} - x_{ij}\beta^{(k)} - \alpha_i^{(k)})^T$$

Equação 31

Passo 7 – Geração segmentada de Σ_y , pois temos σ_1^2 está fixado em 1.

Se σ_1^2 não estivesse fixado em 1, então a condicional completa para Σ_y seria a seguinte:

$$\left[\Sigma_y^{-1(k)} \mid \beta^{(k)}, \alpha^{(k)}, y^{*(k)} \right] = W(A^{(k)}, Nn)$$

$$\Leftrightarrow \left[\Sigma_y^{(k)} \mid \beta^{(k)}, \alpha^{(k)}, y^{*(k)} \right] = IW(A^{(k)}, Nn)$$

Equação 32

No entanto, como $\sigma_1^2 = 1$, é possível gerar Σ_y a partir de $\left[\sigma_{12}^{(k)}, \sigma_{13}^{(k)}, \sigma_{23}^{(k)}, \sigma_2^{2(k)}, \sigma_3^{2(k)} \mid A^{(k)}, \sigma_1^2 = 1 \right]$ dividindo a distribuição Wishart Inversa conforme segue (COWLES, CARLIN e CONNET, 1996):

Primeiro

$$\Sigma_{y22.1}^{(k)} = \begin{bmatrix} \sigma_2^{2(k)} & \sigma_{23}^{(k)} \\ \sigma_{23}^{(k)} & \sigma_3^{2(k)} \end{bmatrix} - \frac{1}{\sigma_1^2} \begin{bmatrix} \sigma_{12}^{(k)} \\ \sigma_{13}^{(k)} \end{bmatrix} \begin{bmatrix} \sigma_{12}^{(k)} & \sigma_{13}^{(k)} \end{bmatrix}$$

Pode ser gerado a partir de

$$\left[\Sigma_{y22.1}^{(k)} \mid A^{(k)}, \sigma_1^2 \right] = IW(A_{22.1}^{(k)}, Nn) \Leftrightarrow \left[\Sigma_y^{-1(k)} \mid A^{(k)}, \sigma_1^2 \right] = W(A_{22.1}^{(k)}, Nn)$$

Equação 33

A outra parte que precisa ser gerada é obtida por meio de

$$\left[\frac{1}{\sigma_1^2} \begin{bmatrix} \sigma_{12}^{(k)} & \sigma_{13}^{(k)} \end{bmatrix} \mid \Sigma_{y22.1}^{(k)}, A^{(k)} \right] = N\left(\frac{1}{a_1^2} \begin{bmatrix} a_{12}^{(k)} & a_{13}^{(k)} \end{bmatrix}, \frac{1}{a_1^2} \Sigma_{y22.1}^{(k)} \right)$$

Equação 34

Feitas essas etapas, os componentes de $\Sigma_y^{(k)}$ são reunidos por meio das seguintes operações:

$$\begin{bmatrix} \sigma_{12}^{(k)} & \sigma_{13}^{(k)} \end{bmatrix} = \frac{1}{\sigma_1^2} \begin{bmatrix} \sigma_{12}^{(k)} & \sigma_{13}^{(k)} \end{bmatrix}, \text{ porque } \sigma_1^2 = 1.$$

e

$$\begin{bmatrix} \sigma_2^{2(k)} & \sigma_{23}^{(k)} \\ \sigma_{23}^{(k)} & \sigma_3^{2(k)} \end{bmatrix} = \Sigma_{y^{22.1}}^{(k)} + \begin{bmatrix} \sigma_{12}^{(k)} \\ \sigma_{13}^{(k)} \end{bmatrix} \begin{bmatrix} \sigma_{12}^{(k)} & \sigma_{13}^{(k)} \end{bmatrix} \quad \text{Equação 35.}$$

3.4 ANÁLISES DE DIAGNÓSTICO DE CONVERGÊNCIA

Uma vez geradas as Cadeias de Markov para cada um dos parâmetros a partir das k iterações, é preciso realizar as análises de diagnóstico para confirmar a convergência dessas cadeias. Embora os algoritmos MCMC permitam uma enorme expansão da classe de modelos possíveis para uma dada base de dados, eles também sofrem de um conhecido problema que pode ser potencialmente sério: geralmente é difícil decidir quando é seguro encerrar as iterações e constatar a convergência da cadeia gerada. Ou seja, a que ponto é razoável acreditar que as amostras são verdadeiramente representativas da distribuição estacionária subjacente a essa Cadeia de Markov? (COWLES e CARLIN, 1996). Para esses autores, essa forma de tratar a convergência é abrangente, porém relevante, pois o que é produzido pelos algoritmos MCMC quando a convergência é constatada é uma amostra de uma distribuição subjacente. Os elementos dessa amostra, porém, devido à natureza dos algoritmos MCMC, serão geralmente correlacionados entre si, o que retarda o objetivo de amostrar valores da distribuição estacionária subjacente.

Diante disso, para se formar uma amostra representativa, é necessário que o processo tenha alcançado convergência para a distribuição estacionária. Para tanto, as b iterações iniciais retiradas da amostra (*burn-in*), bem como as m retiradas entre cada elemento considerado para a amostra final (*thin*), são alternativas indicadas para a obtenção de uma amostra estacionária (BORGES, 2008). Para evidenciar se a amostra atingiu a convergência, são utilizados testes que verificam vieses na amostra que indicam que a mesma não é representativa em relação à distribuição estacionária subjacente e verificam se a cadeia produz estimativas dos parâmetros com variância pequena suficiente para gerar confiança da sua acurácia. A seguir, os testes de diagnóstico de convergência utilizados no presente estudo e aplicados com o auxílio do pacote “*coda*” do *software* R serão apresentados:

3.4.1 Critério de Geweke

Este é um procedimento para teste de convergência sugerido por Geweke (1992), que se utiliza de um teste de igualdade entre as médias da primeira e da última parte da cadeia de Markov. Na literatura, a sugestão é utilizar os primeiros 10% contra os últimos 50% da cadeia

(COWLES e CARLIN, 1996; GEWEKE, 1992). Se a amostra é retirada da distribuição estacionária da cadeia, as duas médias serão iguais, sob a hipótese nula de convergência, e a estatística de teste terá distribuição normal assintótica.

A estatística de teste é um teste Z-score padrão, calculado a partir da diferença entre as duas médias amostrais dividida pelo erro padrão estimado a partir da densidade espectral em zero. O teste, portanto, leva em consideração qualquer autocorrelação existente. Conforme evidenciado por Borges (2008), a estatística de teste é dada por:

$$\frac{\theta_1 - \theta_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \text{Equação 36}$$

Onde θ_1 é a média das n_1 primeiras observações da cadeia, θ_2 é a média das n_2 últimas observações da cadeia e S_1^2 e S_2^2 são os estimadores da variância de θ nas partes inicial e final da cadeia, respectivamente. Admitindo-se que $\frac{n_1}{N}$ e $\frac{n_2}{N}$ são fixos, é possível mostrar que a distribuição da estatística converge para $N(0,1)$ quando $N \rightarrow \infty$. Dessa forma, quando o valor do Z-score estiver dentro do intervalo $[-1,96,1,96]$, que indica o intervalo de confiança de 95% para uma distribuição $N(0,1)$, confirma-se a convergência da cadeia.

3.4.2 Critério de Heidelberger e Welch

Este é um procedimento para teste de convergência sugerido por Heidelberger e Welch (1983), que utiliza a estatística de Cramer-von Mises para testar a hipótese nula de que a amostra de valores de uma determinada variável forma um processo estacionário. Se a hipótese nula for rejeitada ($p\text{-value} < 0,05$) para uma dada variável, o teste deve ser repetido após se descartarem os primeiros 10% das iterações. Se a hipótese nula for novamente rejeitada, o teste deve ser repetido após se descartarem outros 10% das iterações iniciais. Esse processo é repetido até que a hipótese nula seja confirmada, sendo que no máximo 50% das iterações iniciais podem ser descartadas. Se a hipótese nula for novamente rejeitada e já estiver sido descartado o máximo permitido das iterações iniciais, não pode ser confirmada a hipótese nula de estacionariedade e o número de iterações deve ser aumentado. Se a hipótese nula não for rejeitada, então é possível confirmar a estacionariedade da cadeia.

3.4.3 Critério *Half-Width*

Este é um procedimento utilizado para testar se a média estimada está sendo calculada com precisão aceitável. Conforme explicado por Borges (2008), calcula-se um intervalo de confiança com coeficiente de confiança de 95% para a média, utilizando o desvio padrão assintótico por meio da densidade espectral. O resultado passará no teste se a razão entre $(\text{limite superior} - \text{limite inferior})/2$ e a estimativa da média for menor que 0,1.

3.5 CÁLCULO DO *CUSTOMER LIFETIME VALUE* E DO *CUSTOMER EQUITY*

Confirmadas as convergências de cada cadeia de Markov geradas a partir do amostrador de Gibbs para os coeficientes das variáveis independentes (vetores β_1 , β_2 e β_3), bem como para os interceptos específicos de cada cliente (α_{1i} , α_{2i} e α_{3i}), as distribuições posteriores desses parâmetros são, então, obtidas. Conforme Cowles, Carlin e Connet (1996) e Kumar e Shah (2009), um parâmetro é considerado não significativo se um valor zero existir dentro do intervalo entre o percentil 2,5% e o percentil 97,5% da distribuição posterior de cada parâmetro.

Dessa forma, a partir das distribuições posteriores dos parâmetros significantes e dos dados observados para cada uma das variáveis independentes, é possível realizar a previsão da rentabilidade de cada cliente nos 36 meses futuros e, conseqüentemente, calcular o *CLV* de cada cliente e o *CE* dos clientes analisados. A previsão é realizada por meio da distribuição preditiva posterior (GELMAN *et al.*, 2004). Essa distribuição preditiva, cuja forma de amostragem é apresentada a seguir, é calculada a cada período futuro para definir as previsões dos 36 valores futuros para os custos de *marketing*, a probabilidade de compra e a margem de contribuição:

$$\left[y_{new}^{(k)} \mid \alpha_i^{(k)}, \beta^{(k)}, \Sigma_y^{(k)} \right] = N \left(\alpha_i^{(k)} + x_{ij} \beta^{(k)}, \Sigma_y^{(k)} \right) \quad \text{Equação 37}$$

De maneira bastante direta, dados os interceptos de cada regressão específicos para cada cliente ($\alpha_i^{(k)}$), os valores da matriz x_{ij} correspondentes ao período $t - 1$ em relação ao período que está sendo estimado e o vetor de coeficientes ($\beta^{(k)}$) correspondentes às variáveis independentes da matriz x_{ij} , os valores de cada uma das três variáveis determinantes do *CLV* (vetor $y_{new}^{(k)}$) são estimados para cada período t . Como a previsão dos 36 meses futuros se dá

dentro do amostrador de Gibbs, após todos os passos descritos anteriormente para a estimação dos parâmetros, a cada iteração do amostrador de Gibbs, novos valores para os vetores de parâmetros $\alpha_i^{(k)}$ e $\beta^{(k)}$ e para a matriz $\Sigma_y^{(k)}$ são gerados, e a previsão dos 36 meses futuros e do *CLV* para cada cliente é realizada novamente. Ao final das iterações, tendo-se convergência dos parâmetros, obtém-se uma distribuição posterior para o *CLV* de cada cliente, cujas médias são os valores de *CLV* de cada cliente.

Diante disso, dado que os vetores de parâmetros $\alpha_i^{(k)}$ e $\beta^{(k)}$ e a matriz $\Sigma_y^{(k)}$ são gerados a cada iteração, a matriz x_{ij} é o único elemento que precisa ser previamente estimado para os 35 períodos futuros, para que se obtenham os valores estimados para cada variável independente da matriz x_{ij} em cada um dos 35 períodos futuros. Conforme Kumar *et al.* (2008) e Kumar e Shah (2009), os valores foram estimados de 3 maneiras distintas, dependendo da característica da variável independente:

1. Variáveis independentes constantes: Para aquelas variáveis independentes que são constantes para um mesmo cliente ao longo de todos os períodos observados (ex.: frequência de compra), assume-se que o mesmo valor mantém-se nos 35 períodos futuros. Sendo assim, o mesmo valor dos períodos observados é atribuído para os períodos futuros.
2. Variáveis independentes longitudinais:
 - a. Variáveis custos de *marketing* defasado (*lagged*), probabilidade de compra defasada (*lagged*) e margem de contribuição defasada (*lagged*): Conforme Kumar *et al.* (2008), nesses casos, os valores futuros dessas variáveis independentes são estabelecidos a cada estimação do período $t + n$. Dessa forma, os valores previstos para as variáveis dependentes – custos de *marketing*, probabilidade de compra e margem de contribuição – em $t + 1$ são utilizados como variáveis independentes defasadas para prever os valores dessas mesmas variáveis em $t + 2$. Os valores previstos para essas variáveis em $t + 2$ são, por sua vez, utilizados para prever os valores dessas mesmas variáveis em $t + 3$ e, assim, sucessivamente, até os valores previstos em $t + 35$, que são utilizados para prever essas mesmas variáveis em $t + 36$.
 - b. Outras variáveis independentes longitudinais que não são uma das três variáveis dependentes: Nesses casos, como não há uma previsão dessas variáveis sendo realizada a cada período futuro, os 35 valores futuros

das mesmas são estimados separadamente por meio de modelos *ARIMA* (*Autoregressive integrated moving average*) (GUJARATI, 2000). O ajuste dos modelos de cada variável é implementado por meio da função “*auto.arima*” do *software* R, e a previsão dos 35 valores futuros é realizada por meio da função “*forecast*” também do *software* R (ambas as funções são parte do pacote “*forecast*”).

Diante disso, a média e o desvio padrão da função preditiva são definidos para cada período futuro, e os valores para cada uma dessas três variáveis dependentes determinantes do *CLV* podem ser obtidos para cada cliente para cada um dos 36 períodos futuros. Antes de incluir esses valores na fórmula de cálculo do *CLV* (Equação 22), é preciso ainda transformar os valores latentes obtidos para a variável probabilidade de compra de volta aos valores binários 0 ou 1 a partir do seu mapeamento destacado a seguir:

$$\text{Buy}_{ij} = 1, \text{ se } \text{Buy}_{ij}^* > 0$$

$$\text{Buy}_{ij} = 0, \text{ se } \text{Buy}_{ij}^* \leq 0$$

Assim, o *CLV* de cada cliente é calculado a partir da Equação 22. Os fluxos de caixa referentes às margens de contribuição e aos custos de *marketing* são descontados a uma taxa de 1,171% ao mês (equivalente a uma taxa de 15% ao ano) e, então, subtraídos para cada período futuro obtendo-se a rentabilidade mensal prevista de cada cliente. Esses fluxos de caixa resultantes e descontados a valor presente são, então, somados e, assim, o *CLV* de cada cliente é obtido.

O cálculo do *CE*, diante da amostra de 100 clientes considerada no presente estudo, acaba, obviamente, não traduzindo o valor total de toda a base de clientes da empresa. No entanto, diante da facilidade de cálculo do *CE* e do fato de que, caso a base inteira fosse considerada, o cálculo seria o mesmo, optou-se por aplicar o método integralmente e calcular o *CE* sob a suposição de que a amostra de clientes considerada fosse a totalidade da base de clientes da empresa. Sendo assim, o *CE*, por sua vez, é calculado a partir da mensuração de duas parcelas que serão posteriormente somadas para obter o valor do *CE* dinâmico. A primeira parte diz respeito ao *CE* dos clientes retidos na base (CE_R), que é calculado a partir do somatório dos *CLVs* obtidos para cada um dos clientes:

$$CE_R = \sum_{i=1}^N CLV_i \quad \text{Equação 38}$$

A outra parcela para o cálculo do CE diz respeito à previsão do valor dos clientes que a firma espera adquirir no futuro (CE_A). Para o cálculo do CE_A , optou-se por adotar o mesmo método utilizado por Kumar e Shah (2009). Para tanto, parte-se de um percentual de crescimento anual da base de clientes de acordo com o plano de crescimento da firma (ex.: 3%) para a obtenção do número M de clientes que serão adquiridos nos próximos três anos (36 meses). A partir disso, a seguinte equação é aplicada:

$$CE_A = \sum_{K=1}^M \sum_{j=T+1}^{T+36} \frac{CLV_{kj} - A_{kj}}{(1+r)^{j-T}} \quad \text{Equação 39}$$

Onde

CLV_{kj} = O valor vitalício para o cliente k que se espera que seja adquirido no tempo j ;

A_{kj} = O custo de aquisição correspondente ao cliente k que se espera que seja adquirido no tempo j ;

j = índice para os períodos de tempo, em meses, no caso do presente estudo;

T = Demarcação do final do tempo total de calibração ou observação;

r = Taxa de desconto, mensal no caso do presente estudo;

O CLV_{kj} é igual à média dos $CLVs$ dos clientes retidos na base da firma. A_{kj} , por sua vez, é igual ao custo de aquisição médio considerado pela firma.

Diante disso, o valor total do CE dinâmico é calculado a partir da seguinte equação:

$$CE = CE_R + CE_A \quad \text{Equação 40}$$

3.6 SEGMENTAÇÃO DA BASE DE CLIENTES A PARTIR DO *CUSTOMER LIFETIME VALUE*

Uma vez calculados os $CLVs$ de cada um dos clientes, serão calculados os decis referentes à distribuição do valor adicionado por cada cliente retido na base ao valor do CE_R . Cada decil representará um segmento distinto. No entanto, caso seja necessário, os decis que representarem a maior parcela de valor adicionado podem ser unidos formando o segmento de clientes com alto CLV , os decis intermediários podem ser unidos formando o segmento de clientes com CLV médio, ao passo que os decis com a menor parcela de valor adicionado ou

com valor negativo podem ser unidos formando o segmento de clientes com baixo/negativo *CLV*.

3.7 VERIFICAÇÃO DA CAPACIDADE DE PREVISÃO DO MODELO APLICADO

A cada possível modelo testado, 3 testes foram realizados para verificar a sua capacidade de previsão:

1. Conforme Venkatesan, Kumar e Bohling (2007), as previsões de *CLV* realizadas a partir do modelo aplicado foram comparadas com aquelas realizadas por um modelo denominado “ingênuo”. O *CLV* calculado pelo modelo ingênuo é obtido a partir apenas das médias dos custos de *marketing* e das margens de contribuição de cada cliente, que são repetidas sucessivamente formando os 36 fluxos de caixa futuros, que descontados a valor presente dão origem ao *CLV*.
2. Os valores de *CLV* previstos para cada cliente foram analisados em relação ao histórico de transações desses clientes, buscando confirmar a consistência das previsões realizadas. Essa verificação foi possível devido à reduzida quantidade de clientes analisados.
3. Conforme Kumar *et al.* (2008), Venkatesan, Kumar e Bohling (2007) e diante da limitada quantidade de períodos disponíveis na amostra, optou-se por comparar os valores previstos em relação aos valores reais, observados, para o período $t + 1$. Para tanto, o último período observado foi retirado da base de dados e o modelo foi novamente calculado. Ao final, os valores previstos das três variáveis dependentes para o período $t + 1$ foram comparados aos valores reais.

A partir desses testes, busca-se verificar a consistência das estimações realizadas para cada cliente. Como o objetivo do modelo aplicado é estimar de maneira adequada o valor futuro de cada cliente é indispensável checar se o *CLV* de cada cliente realmente traduz o que pode ser esperado para cada cliente diante de seu histórico de transações realizadas, bem como checar se os resultados obtidos são melhores do que se uma forma determinística de estimação do *CLV* fosse empregada.

4 RESULTADOS

Diante do método estabelecido, os procedimentos adequados para a aplicação do modelo e para a realização dos testes propostos foram realizados com o auxílio do *software* R. Sendo assim, seguindo o método proposto, a aplicação do modelo seguiu as seguintes etapas:

1. Análises descritivas;
2. Estimação dos parâmetros;
3. Cálculo do *CLV* e do *CE* da amostra considerada;
4. Segmentação da amostra de clientes com base no *CLV*;
5. Análise da capacidade de previsão.

Nesta seção, os resultados obtidos em cada uma dessas etapas serão apresentados.

4.1 ANÁLISES DESCRITIVAS

Com o intuito de conhecer as informações existentes, bem com a relação entre as variáveis consideradas, análises descritivas simples foram conduzidas. A seguir, os principais resultados desta etapa serão apresentados.

4.1.1 Análises Univariadas

As primeiras análises realizadas foram as estatísticas básicas de cada variável. Os resultados das Tabelas 4 e 5 auxiliam na tarefa de compreender o comportamento de cada variável com escala de razão e de cada variável categórica, respectivamente.

Tabela 4 – Análises Univariadas – variáveis de razão

	Número de casos		Média	Mediana	Desvio Padrão	Mínimo	Máximo	Percentis		
	Válidos	Missing						25%	50%	75%
MC total	855	1846	251,67	20,21	837,69	0,00	9.018,79	4,17	20,21	100,89
Custos de marketing	2700	0	3,76	0,00	6,43	0,00	28,15	0,00	0,00	5,91
Cross-buying	2700	0	0,38	0,00	0,61	0,00	4,00	0,00	0,00	1,00
Total de dias com transação	2700	0	1,47	0,00	4,47	0,00	33,00	0,00	0,00	0,00
Períodos com compras	100	0	9,10	7,00	7,71	1,00	28,00	3,00	7,00	15,50
Net médio	100	0	45.191,82	6.926,70	171.083,86	30,31	1.616.045,87	1.255,09	6.926,70	19.394,92
Share-of-wallet	100	0	25,85	14,83	25,05	0,16	96,93	6,67	14,83	39,30
Frequência	100	0	5,75	2,00	7,62	1,00	29,00	1,00	2,00	6,00
Recência	100	0	6,81	3,00	8,41	0,00	27,00	0,00	3,00	10,00
Salário	100	0	8.551,71	5.879,50	9.717,37	500,00	50.000,00	2.882,63	5.879,50	10.000,00
Aplicações	100	0	153.240,93	50.000,00	276.870,97	500,00	2.000.000,00	10.000,00	50.000,00	165.000,00
Idade	100	0	41,00	37,50	12,54	24,00	82,00	31,25	37,50	48,75

Diante desses resultados, é possível conhecer o comportamento de cada variável. A margem de contribuição total apresenta uma média de R\$ 251,67 e valor máximo de R\$ 9.018,79, sendo que o percentil 75% possui valor R\$ 100,89. Diante disso, é possível concluir que a distribuição da margem de contribuição observada é bastante concentrada em valores abaixo de R\$ 100,89. Além disso, é possível verificar a grande quantidade de *missing values* existente na base de dados, pois, geralmente, os clientes não compram todos os meses do ano.

A análise univariada da variável *cross-buying*, que indica a quantidade de categorias de produtos adquiridas pelo cliente em um dado mês, mostra que os clientes, quando compram, geralmente compram apenas uma categoria de produtos por mês. Os resultados para a variável custos de *marketing* confirmam o que foi afirmado pelos profissionais da empresa: “pouco investimento em incentivos de *marketing*”. Diante desse investimento restrito, a média mensal de custos de *marketing* por cliente é de R\$ 3,76, sendo que o maior valor mensal gasto com contato de *marketing* foi de R\$ 28,15.

O número de períodos (meses) com compra evidencia que, em média, os clientes realizam compras em apenas 9 dos 29 meses inicialmente observados. Em relação a essa variável, ressalva deve ser feita para o fato de que cerca 52% dos clientes presentes na amostra iniciaram sua relação com a empresa durante o período de análise. Para que fosse levada em consideração essa característica, o cálculo da frequência de compra foi realizado com base no total de meses que o cliente possui de relacionamento com a empresa (clientes anteriores a janeiro de 2011 possuem 29 meses no total, enquanto que, para clientes posteriores a janeiro de 2011, esse número base é calculado somente a partir do mês em que seu cadastro foi realizado). Dessa forma, a frequência de compra da amostra analisada apresenta uma média de 1 compra a cada 5,75 meses. A média da variável recência de compra, por sua vez, evidencia que em média a última compra de cada cliente foi a 6,81 meses.

A variável *net médio* indica que os clientes possuem em média uma posição financeira de R\$ 45.191,82, no entanto essa média é fortemente influenciada por um único cliente que possui *net médio* de R\$ 1.616.045,87. Como a amostragem foi aleatória, a inclusão ou não desse cliente não pode ser controlada, no entanto, embora as suas margens de contribuição estejam entre as maiores, elas não destoam das outras margens de contribuição na mesma magnitude que ocorre com a variável *net médio*. Se observarmos o percentil 75% da variável *net médio* (R\$ 19.394,92), podemos constatar que há uma concentração dos valores de *net médio* abaixo desse patamar. As aplicações totais declaradas pelos clientes também possuem

uma média 153.240,93, que é influenciada agora por três clientes que possuem aplicações em com total superior a R\$ 1.000.000,00. O percentil 75% da variável aplicações (R\$ 165.000,00) evidencia que há uma concentração dos valores de aplicações abaixo desse patamar.

Por fim, os salários dos clientes analisados possuem média de R\$ 8.551,71 e os clientes possuem idade média de 41 anos, sendo que o investidor mais jovem possui 24 anos e o mais velho possui 82 anos.

Tabela 5 – Análises Univariadas – variáveis de categóricas

Sexo		Frequência	Perfil de risco		Frequência
Número de casos	Feminino	28	Número de casos	Conservador	22
	Masculino	72		Arrojado	78
	Total	100		Total	100
Estado Civil		Frequência	Tempo de relacionamento		Frequência
Número de casos	Casado	34	Número de casos	Posterior a jan 2011	52
	Solteiro	66		Anterior a jan 2011	48
	Total	100		Total	100
Categoria de produtos majoritária		Frequência	Tipo de cliente		Frequência
Número de casos	BMF	20	Número de casos	B2C	7
	Bovespa	58		B2B	93
	Fundos	18		Total	100
	Renda Fixa	4			
	Total	100			

As análises das variáveis categóricas são utilizadas para caracterização da amostra analisada. De forma bastante direta, é possível identificar que 72% da amostra é composta por homens. 66% da amostra é composta por clientes solteiros. 58% dos clientes investem majoritariamente na categoria de produtos Bovespa, 20% investem majoritariamente na categoria de produtos BMF e apenas 22% investem majoritariamente ou em Fundos ou em Renda Fixa. 52% dos clientes entraram na base de dados após janeiro de 2011. E, por fim, 93% dos clientes são do tipo B2B, o que é bastante coerente visto a grande quantidade de escritórios afiliados que a empresa possui frente à capacidade de atendimento dos escritórios controlados pela matriz.

4.1.2 Análises de Correlação

Tendo como objetivo conhecer as correlações significativas e mais altas entre as variáveis candidatas a preditoras e as variáveis dependentes, as correlações de Pearson foram calculadas e os resultados estão evidenciados na Tabela 6, ordenados pela força da correlação entre cada variável candidata a preditora e cada variável dependente.

É importante destacar que, ao analisarem-se as correlações entre as variáveis dependentes, constata-se que os custos de *marketing* possuem uma alta correlação com a variável dependente *buy* ($,809^{**}$), o que seria previsível, pois os custos de *marketing* foram apropriados *a posteriori* a partir da incidência de compra e da posição financeira de cada cliente (*net*), diante da política de relacionamento com clientes da empresa. Em contrapartida, a correlação entre os custos de *marketing* e a variável dependente MC Total não é significativa a um nível de 0,05. O cálculo da correlação entre as variáveis *Buy* e MC Total não pode ser realizado sem que os *missings values* da margem de contribuição sejam tratados. Como a incidência de compra será constante com valor 1 quando a variável MC total for diferente de zero, pois quando a incidência de compra é igual a zero, a variável MC total é *missing value*.

O cálculo das correlações entre as três variáveis dependentes e as variáveis candidatas a preditoras foi de grande importância para a obtenção do modelo final calculado. A partir do conhecimento dessas correlações, os modelos testados foram construídos com maior segurança e, assim, o caminho para a definição do modelo com melhor ajuste pôde ser mais direto.

4.2 ESTIMAÇÃO DOS PARÂMETROS

A partir do método descrito na subseção 3.2, o modelo foi estimado para a amostra selecionada por meio do amostrador de Gibbs. O tempo de processamento, devido à já citada necessidade de otimização do código criado no *software* R, foi o grande limitador para a aplicação do modelo para toda a base de dados. Diante de pareceres de *experts* da área de tecnologia da informação que trabalham com o *software* R, uma vez criado o código que atenda corretamente às operações matemáticas necessárias, é necessário um novo trabalho de programação para tornar as operações que estão sendo realizadas por laços (*for*) em operações entre vetores. As operações entre vetores geram elevados ganhos de processamento, principalmente quando o volume de dados aumenta. Essa indicação foi seguida em algumas partes do código e os efeitos foram bastante positivos, no entanto os laços (*for*) que custam mais tempo de processamento ainda permanecem.

Tabela 6 – Análises de correlação

Variáveis candidatas a predictoras	Custos de marketing			Variáveis candidatas a predictoras	Buy			Variáveis candidatas a predictoras	MC Total		
	Correlação de Pearson	Sig. (2-tailed)	N		Correlação de Pearson	Sig. (2-tailed)	N		Correlação de Pearson	Sig. (2-tailed)	N
Buy	,819**	0,000	2700	Cross-buying	,904**	0,000	2700	MC Total lag 1	,536**	,000	855
Cross-buying	,743**	0,000	2700	Custos de marketing	,819**	0,000	2700	Total de dias com transação	,505**	,000	855
Buy lag 1	,555**	,000	2700	Buy lag 1	,676**	0,000	2700	MC Total lag 2	,424**	,000	855
Custos de marketing lag 1	,539**	,000	2700	Buy lag 2	,592**	,000	2700	BMF é categoria majoritária	,336**	,000	855
Buy lag 2	,509**	,000	2700	Períodos com compras	,592**	,000	2700	Net	,288**	,000	855
Cross-buying anual	,490**	,000	2700	Cross-buying anual	,562**	,000	2700	Cross-buying	,231**	,000	855
Períodos com compras	,475**	,000	2700	Custos de marketing lag 1	,546**	,000	2700	Net médio anual	,230**	,000	855
Custos de marketing lag 2	,466**	,000	2700	Total de dias com transação	,481**	,000	2700	Net médio	,229**	,000	855
Total de dias com transação	,426**	,000	2700	Custos de marketing lag 2	,462**	,000	2700	Cross-buying anual	,215**	,000	855
Frequência de compra	-,268**	,000	2700	Frequência de compra	-,321**	,000	2700	Aplicações	,189**	,000	855
Recência de compra	-,231**	,000	2700	Recência de compra	-,272**	,000	2700	Salário	,184**	,000	855
Net	,214**	,000	2700	Share-of-wallet	,267**	,000	2700	Fundos é categoria majoritária	-,156**	,000	855
Share-of-wallet	,211**	,000	2700	Net	,233**	,000	2700	Bovespa é categoria majoritária	-,145**	,000	855
MC Total lag 1	,196**	,000	2700	Net médio anual	,205**	,000	2700	Estado Civil (1=solteiro)	-,139**	,000	855
MC Total lag 2	,182**	,000	2700	MC Total lag 1	,202**	,000	2700	Perfil de risco	,138**	,000	855
Net médio anual	,181**	,000	2700	MC Total lag 2	,173**	,000	2700	Tipo de cliente (1=B2B)	,091**	,008	855
Net médio	,165**	,000	2700	Net médio	,173**	,000	2700	Idade	,087**	,011	855
Aplicações	,139**	,000	2700	Aplicações	,140**	,000	2700	Custos de marketing lag 1	,073**	,033	855
Renda Fixa é categoria majoritária	-,058**	,003	2700	Tipo de cliente (1=B2B)	-,110**	,000	2700	Buy lag 1	,073*	,033	855
Fundos é categoria majoritária	,053**	,005	2700	Fundos é categoria majoritária	,098**	,000	2700	Recência de compra	,069**	,044	855
Salário	,045**	,020	2700	Renda Fixa é categoria majoritária	-,078**	,000	2700	Buy lag 2	,062**	,070	855
Tipo de cliente (1=B2B)	,042**	,030	2700	Salário	,066**	,001	2700	Custos de marketing lag 2	,059**	,086	855
MC Total	,031**	,360	855	Perfil de risco	-,054**	,005	2700	Renda Fixa é categoria majoritária	,048**	,158	855
Estado Civil (1=solteiro)	,023**	,225	2700	Idade	,040**	,039	2700	Tempo de relacionamento	-,041**	,230	855
Idade	,022**	,243	2700	Bovespa é categoria majoritária	-,030**	,113	2700	Custos de marketing	,031**	,360	855
Perfil de risco	-,022**	,250	2700	Tempo de relacionamento	,023**	,227	2700	Períodos com compras	-,029**	,399	855
Bovespa é categoria majoritária	-,017**	,388	2700	BMF é categoria majoritária	-,018**	,352	2700	Frequência de compra	-,018**	,608	855
Sexo	,014**	,461	2700	Estado Civil (1=solteiro)	,014**	,467	2700	Sexo	,016**	,640	855
Tempo de relacionamento	,005**	,791	2700	Sexo	,006**	,754	2700	Share-of-wallet	-,005**	,881	855
BMF é categoria majoritária	-,002**	,900	2700	Buy	1		2700	MC Total	1		855
Custos de marketing	1		2700	MC Total	^a	0,000	855	Buy	^a	0,000	855

** Correlação é significativa a 0.01 (2-tailed).

* Correlação é significativa a 0.05 (2-tailed).

a. Não pode ser computado, porque pelo menos uma variável é constante.

Uma vez que essa etapa chamada de “vetorização” seja cumprida, os pareceres recebidos dos *experts* da área de tecnologia da informação indicam que o tempo de processamento pode ser melhorado ainda mais por meio da utilização de processamento paralelo em dois ou mais núcleos de processamento do computador utilizado. O *software* R foi construído de forma que apenas um núcleo de processamento do computador seja utilizado, sendo que, atualmente, computadores comuns possuem mais de um núcleo de processamento. Isso implica a subutilização da capacidade de processamento do computador. A melhoria do tempo de processamento, portanto, poderá ser obtida por meio de um novo esforço de programação para tornar o código “vetorizado” capaz de utilizar dois ou mais núcleos de processamento, dependendo da quantidade de núcleo de processamento existente na máquina utilizada.

Sendo assim, embora o modelo aplicado necessite de centenas ou até milhares de iterações que envolvem operações custosas em termos de processamento, existem alternativas mais eficientes sob o prisma de velocidade de processamento, que podem ser implementadas para reduzir substancialmente os tempos de processamento e tornar o modelo aplicável para grandes bases de dados. Diante disso, essa dificuldade encontrada na presente aplicação do modelo não é de maneira alguma um impeditivo para a aplicação em larga escala do modelo construído.

Cada iteração do amostrador de Gibbs apresentou um tempo médio de processamento de 10,5 segundos. Se a previsão do *CLV* de cada cliente é incluída dentro do laço do amostrador Gibbs, então o tempo médio de processamento de cada iteração sobe para 35 segundos. Para o modelo final aplicado, foram necessárias 700 iterações, com um *burn-in* de 100 iterações e sem *thinning* para alcançar a convergência dos parâmetros. Caso um número maior de clientes fosse utilizado, os tempos de processamento, devido à pouca eficiência computacional do código utilizado, seriam exponencialmente elevados, o que impossibilitaria a aplicação do modelo para o presente trabalho, devido à elevada quantidade de modelos alternativos que precisaram ser testados.

Diante da amostra aleatoriamente definida de 100 clientes, as variáveis candidatas a variável independente foram testadas e as variáveis apresentadas na Tabela 7, acompanhadas das médias e desvios padrão de suas respectivas distribuições posteriores, foram definidas como as variáveis preditoras do modelo aplicado:

Tabela 7 – Parâmetros estimados

Variável Independente	Média	Desvio Padrão
Regressão 1: Custos de <i>marketing</i>		
Custos de <i>marketing</i> Lag 1	0,059	0,044
Custos de <i>marketing</i> Lag 2	0,255	0,020
MC Total Lag 1	0,187	0,006
<i>Net</i>	0,024	0,023
Regressão 2: Incidência de compra		
<i>Cross-buying</i>	0,504	0,012
<i>Buy</i> Lag 1	0,146	0,021
<i>Buy</i> Lag 2	0,100	0,015
Frequência de compra	-0,090	0,001
Regressão 3: Margem de contribuição		
<i>Cross-buying</i>	0,031	0,005
MC Total Lag 1	0,129	0,003
MC Total Lag 2	0,021	0,002
Total de dias com transação	0,293	0,006
<i>Net</i>	0,192	0,011

Entre as três regressões, aquela referente aos custos de *marketing* foi a que apresentou maior dificuldade de definição de suas variáveis preditoras e, conforme será apresentado na seção de verificação da capacidade de predição do modelo, a previsão poderia ainda ter sido mais precisa. Essa questão, porém, se deu devido principalmente ao fato de os custos de *marketing* terem sido apropriados para cada cliente de acordo com a política de relacionamento com os clientes definida pela empresa. Com isso, os custos de *marketing* dependem principalmente da incidência ou não de compra e da posição financeira do cliente (*net*). Os modelos testados apenas com a variável *Buy lag* 1 não tiveram um bom ajuste, e as previsões acabaram ficando bastante inconsistentes. Dessa forma, optou-se pelo conjunto de variáveis apresentado na Tabela 7, pois foi a alternativa que possibilitou a maior consistência nos valores previstos de custos de *marketing* para cada cliente. Assim, os custos de *marketing* são previstos em função dos valores dos custos de *marketing* nos dois períodos anteriores, bem como do valor da margem de contribuição no período anterior e o valor da posição financeira do cliente.

Em relação à regressão 2, a quantidade de categorias de produtos diferentes transacionados pelos clientes apresentou alto grau de associação com a variável dependente

incidência de compra. Além da variável *cross-buying*, a própria incidência ou não de compra nos dois períodos anteriores e a frequência de compra do cliente também compõem o conjunto de variáveis preditoras da probabilidade de o cliente comprar, sendo que a frequência de compra possui relação inversa com a incidência de compra, pois, quanto mais o cliente transaciona ao longo do período analisado, menor é a sua frequência de compra.

Por fim, na regressão 3, a variável total de dias com transação apresentou importante nível de associação com o valor da margem de contribuição gerada por cada cliente. Isso ocorre, pois os clientes que investem nas categorias de produtos BMF e Bovespa são aqueles que geram maior margem de contribuição. A posição financeira do cliente, conforme já esperado, também possui uma associação positiva com a margem de contribuição. Além dessas duas variáveis, os valores da margem de contribuição nos dois períodos anteriores e a variável *cross-buying* também compõem o conjunto de variáveis preditoras da regressão 3.

4.3.1 Análises de diagnóstico de convergência

A partir dos critérios apresentados na seção 3, a convergência dos parâmetros pôde ser verificada. A Tabela 8 apresenta os valores dos critérios de Geweke, de Heidelberger e Welch e de *Halfwidth* para as distribuições posteriores dos coeficientes (β) das variáveis independentes.

Tabela 8 – Resultados para os testes de convergência

	Geweke	Heidelberger e Welch	<i>Halfwidth</i>
Variável independente	<i>Z-Score</i>	<i>p-value</i>	-
Custos de <i>marketing Lag 1</i>	-1,356	0,553	0,006
Custos de <i>marketing Lag 2</i>	-1,155	0,417	0,005
MC Total <i>Lag 1</i>	0,461	0,159	0,000
<i>Net</i>	1,239	0,715	0,006
<i>Cross-buying</i>	1,134	0,654	0,002
<i>Buy Lag 1</i>	-1,082	0,532	0,004
<i>Buy Lag 2</i>	-0,968	0,672	0,003
Frequência de compra	0,007	0,978	0,000
<i>Cross-buying</i>	0,985	0,158	0,000
MC Total <i>Lag 1</i>	1,105	0,205	0,000
MC Total <i>Lag 2</i>	0,872	0,410	0,000
Total de dias com transação	-0,961	0,181	0,000
<i>Net</i>	-1,308	0,695	0,003

Os valores dos *Z-scores* do teste de Geweke estão dentro do intervalo $[-1,96, 1,96]$ para todas as variáveis independentes, o que indica igualdade entre as médias da primeira (10%) e da última parte (50%) das cadeias de Markov formadas. Os *p-values* do teste de Heidelberger e Welch estão todos acima de 0,05, o que indica que a hipótese nula de estacionariedade das cadeias de Markov formadas não foi rejeitada. O teste *Halfwidth*, por sua vez, apresentou valores abaixo de 0,1 para todas as cadeias de Markov formadas, o que assegura que a média calculada para cada cadeia possui precisão aceitável.

Diante desses resultados, é possível confirmar que as cadeias formadas após todas as iterações do amostrador de Gibbs para cada variável independente apresentam convergência e, portanto, essas distribuições posteriores formadas podem ser consideradas amostras de suas verdadeiras distribuições posteriores subjacentes, respectivamente.

Além dos três testes apresentados, é também uma boa prática checar a densidade de cada distribuição posterior, bem como quais foram os valores amostrados em cada iteração do amostrador de Gibbs. Essas informações estão representadas nas Figuras a seguir pelo gráfico de densidade de cada distribuição posterior e pelo gráfico de traço de cada cadeia formada, respectivamente.

Figura 5 – Gráficos de densidade e traço da cadeia para a variável custos de *marketing lag 1* – Regressão 1

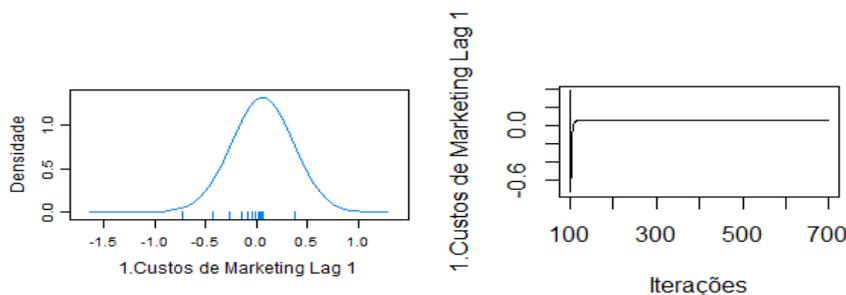


Figura 6 – Gráficos de densidade e traço da cadeia para a variável custos de *marketing lag 2* – Regressão 1

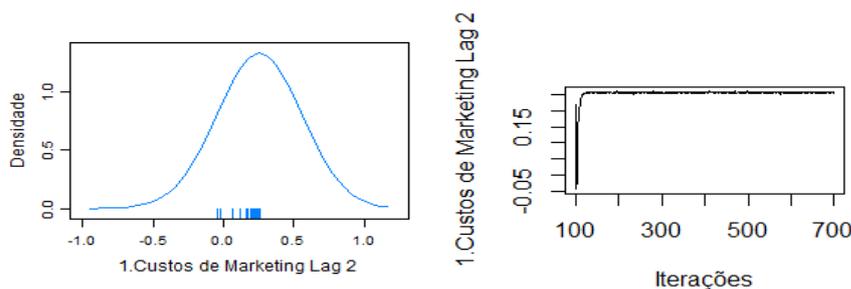


Figura 7 – Gráficos de densidade e traço da cadeia para a variável MC total lag 1 – Regressão 1

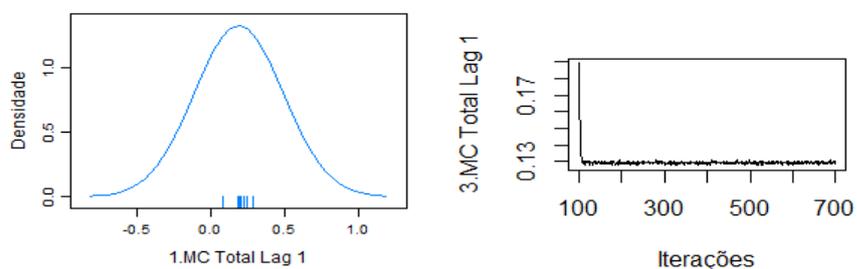


Figura 8 – Gráficos de densidade e traço da cadeia para a variável net – Regressão 1

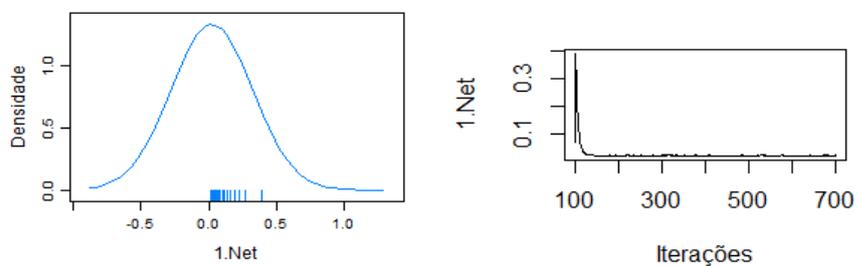


Figura 9 – Gráficos de densidade e traço da cadeia para a variável cross-buying – Regressão 2

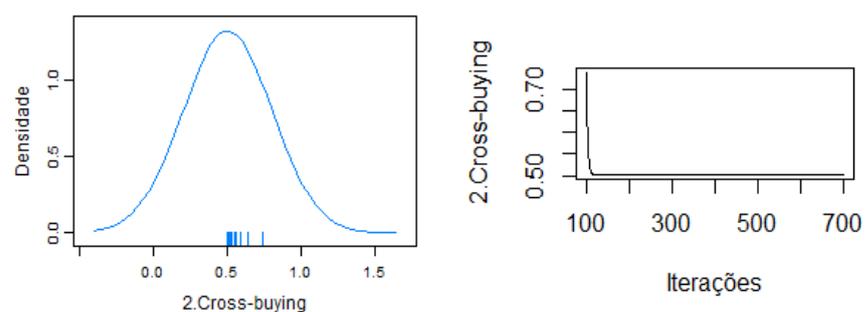


Figura 10 – Gráficos de densidade e traço da cadeia para a variável buy lag1 – Regressão 2

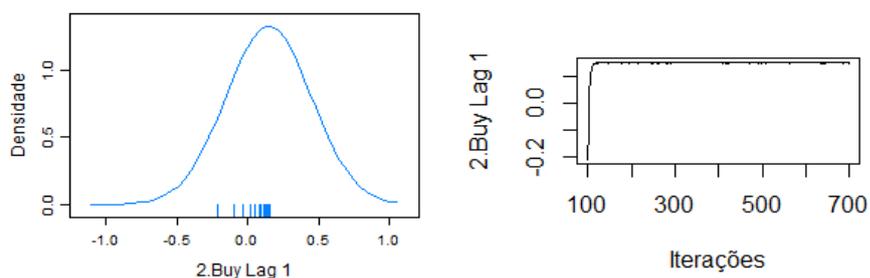


Figura 11 – Gráficos de densidade e traço da cadeia para a variável *buy lag2* – Regressão 2

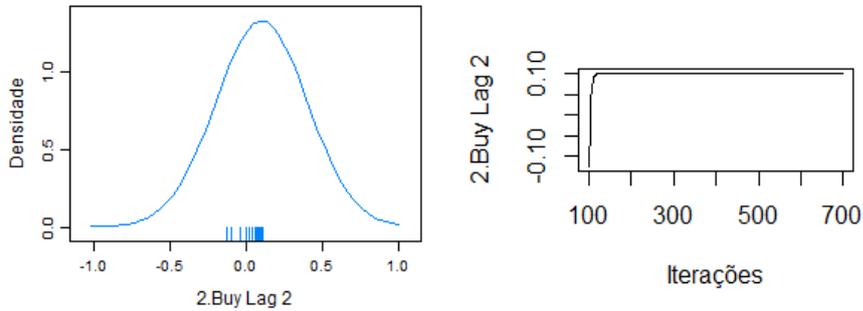


Figura 12 – Gráficos de densidade e traço da cadeia para a variável frequência de compra – Regressão 2

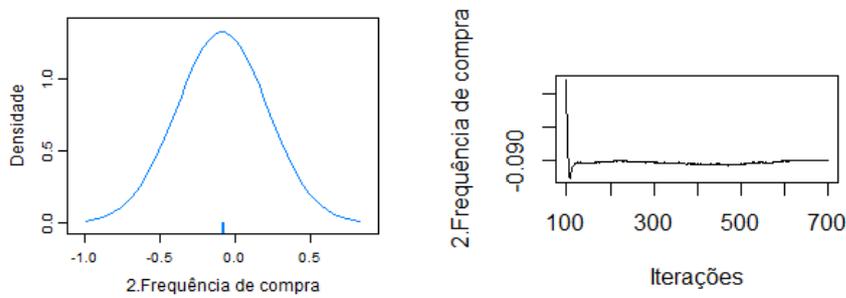


Figura 13 – Gráficos de densidade e traço da cadeia para a variável *cross-buying* – Regressão 3

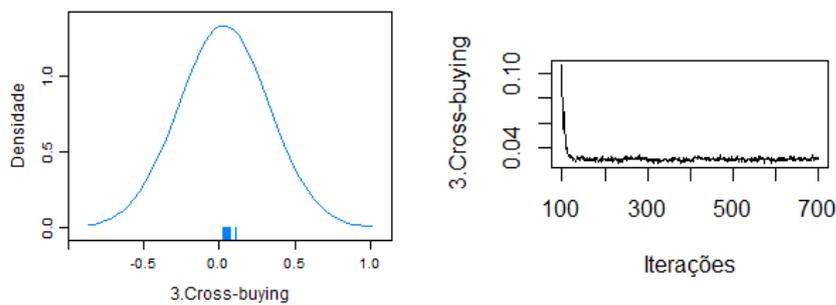


Figura 14 – Gráficos de densidade e traço da cadeia para a variável MC total lag 1 – Regressão 3

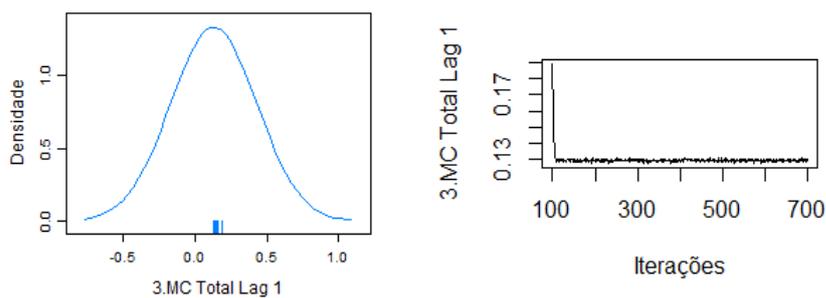


Figura 15 – Gráficos de densidade e traço da cadeia para a variável MC total lag 2 – Regressão 3

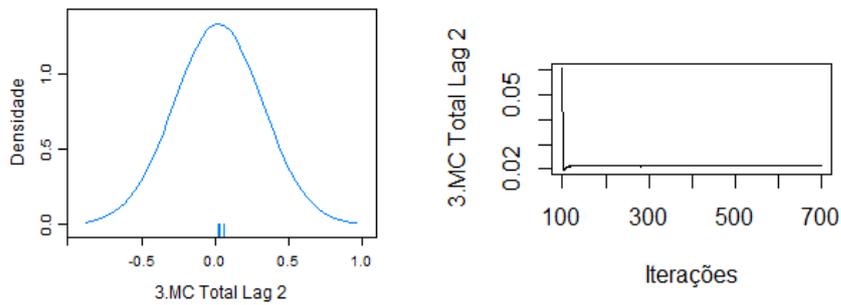


Figura 16 – Gráficos de densidade e traço da cadeia para a variável total de dias com transação – Regressão 3

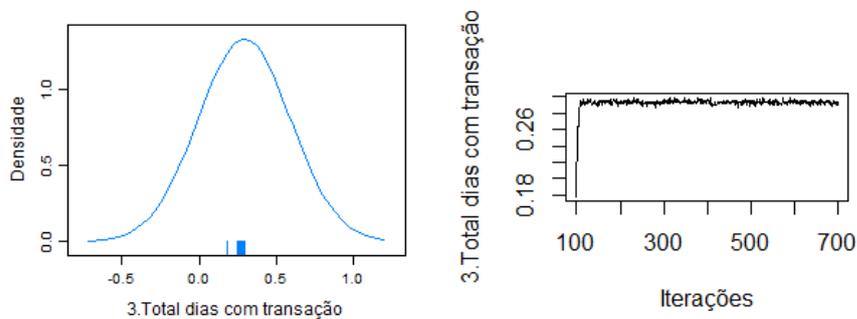
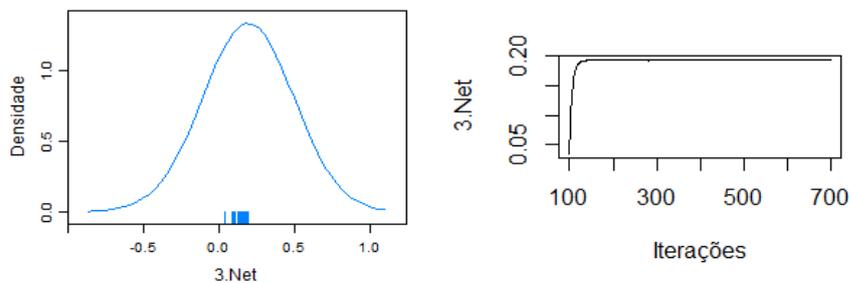


Figura 17 – Gráficos de densidade e traço da cadeia para a variável net – Regressão 3



Os gráficos apresentados reforçam a condição de convergência das cadeias de Markov formadas. A análise dos gráficos de densidade é importante para constatar que os valores amostrados geram uma distribuição posterior, cujos valores amostrados possuem maior densidade de valores próximos ao valor médio dessa distribuição. Já a análise dos gráficos de traço das cadeias é importante para constatar que, à medida que cada iteração é feita, os valores amostrados convergem para um valor comum em todas as iterações posteriores.

As análises de convergência a partir dos critérios de Geweke, de Heidelberger e Welch e de *Halfwidth* para os 300 interceptos, 100 para cada regressão, são apresentados no Anexo A. Analisando os *Z-scores* do teste de Geweke, é possível verificar que a grande maioria dos

interceptos se encontram dentro do intervalo $[-1,96, 1,96]$ e um número pequeno de interceptos apresentaram valores fora desse intervalo, porém não foram valores extremos. Quando os poucos interceptos que não atendem ao critério de Geweke, bem como todos os outros interceptos são analisados em relação tanto ao critério de Heidelberger e Welch quanto ao critério de *Halfwidth*, pôde-se constatar que todos esses interceptos atendem aos critérios desses testes. Sendo assim, diante dos resultados positivos desses dois últimos testes e da confirmação da estacionariedade de todas as cadeias formadas, assume-se também a convergência das cadeias dos interceptos. As médias e os desvios padrão das distribuições posteriores de cada intercepto são apresentadas no Anexo B.

4.3 CÁLCULO DO *CUSTOMER LIFETIME VALUE* E DO *CUSTOMER EQUITY*

Uma vez estimados os parâmetros, é possível realizar a previsão dos 36 fluxos de caixa futuros para cada cliente. De forma direta, os fluxos de caixa para os custos de *marketing* são os próprios valores previstos. Os fluxos de caixa para a margem de contribuição são determinados pela multiplicação entre o valor previsto para a margem de contribuição e o valor previsto que indica se o cliente comprará ou não. Este último valor é determinado a partir do valor previsto para a variável latente *Buy**. Caso esse valor seja menor ou igual a 0, então o cliente não comprará, e o valor previsto para a margem de contribuição é multiplicado por 0. Agora, caso o valor previsto para a variável latente *Buy** seja maior do que 0, então o cliente não comprará, e o valor previsto para a margem de contribuição é multiplicado por 1.

Com os 36 fluxos de caixa futuros para cada cliente definidos, a próxima tarefa é trazê-los a valor presente, conforme Equação 22 e, após, somar os fluxos de caixa resultantes para obter o valor de *CLV* de cada cliente. Nesse ponto, cabe ressalva em relação ao método adotado, pois a previsão dos 36 meses é realizada a partir dos 29 períodos observados disponíveis. É esperado, portanto, uma maior fragilidade das previsões realizadas, no entanto a base de dados utilizada possui apenas 29 meses disponíveis com dados consistentes e realizar a estimação do valor vitalício de um cliente com menos de 36 meses futuros contrariaria a literatura existente sobre o tema. Diante disso, buscou-se realizar o estudo conciliando o método com os recursos informacionais disponíveis. A Tabela 9, presente na seção de verificação da capacidade de previsão do modelo, traz o valor de *CLV* estimado para cada um dos 100 clientes analisados.

Uma vez estimado o *CLV*, é possível estimar também o *CE*. O *CE*, em seu conceito, refere-se ao valor estimado de toda a base de clientes de uma empresa. Conforme já mencionado, no presente estudo, apenas uma amostra de 100 clientes foi utilizada, o que faz com que o *CE* real da empresa não possa ser estimado. No entanto visando a testar o modelo construído, bem como aplicar o método integralmente, decidiu-se estimar o valor da amostra de 100 clientes. Claramente, esse valor de *CE* calculado não é o valor real da empresa, mas, sim, uma suposição de que, caso a empresa tivesse apenas 100 clientes, esse seria o valor de sua base de clientes. Sendo assim, esse valor é calculado a partir da soma de duas parcelas: o *CE* dos clientes retidos, que é simplesmente a soma dos *CLVs* estimados para cada cliente, e o *CE* dos clientes que serão adquiridos, que é calculado conforme a Equação 39. Para o cálculo dessa segunda parcela, assumiu-se o objetivo de crescimento da base de clientes de 12% ao ano e um custo médio de aquisição de cada cliente de R\$ 112,60, o que representa 4 vezes o maior valor mensal gasto para reter um cliente. O valor do *CLV* médio dos clientes retidos é R\$ 2054,63. Dessa forma, o valor do *CE*, caso a empresa possuísse apenas os 100 clientes analisados, seria o seguinte:

$$CE_R = \text{R\$ } 205.463,20$$

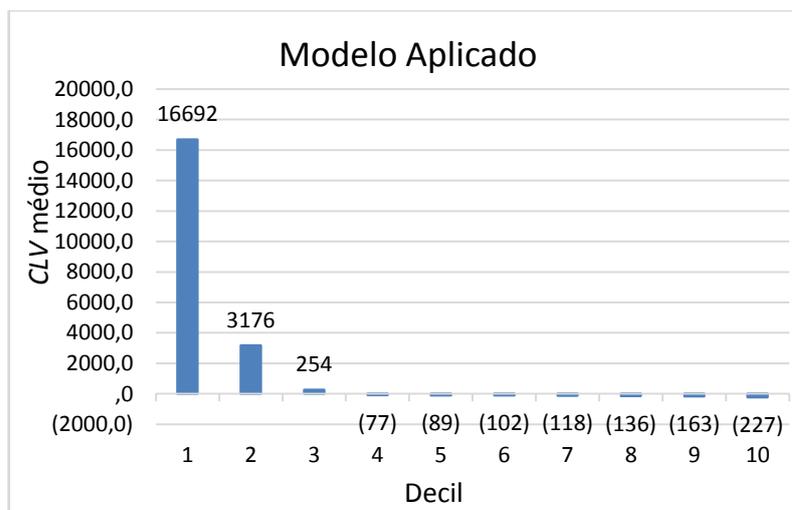
$$CE_A = \text{R\$ } 56.774,99$$

$$CE_{100} = \text{R\$ } 205.463,20 + \text{R\$ } 56.774,99 = \text{R\$ } 262.238,19$$

4.4 SEGMENTAÇÃO DA AMOSTRA DE CLIENTES A PARTIR DO *CLV*

Diante do cálculo do *CLV* de cada cliente, o último objetivo específico proposto para o presente estudo pôde ser cumprido. A segmentação de clientes é uma importante ferramenta de gestão para direcionar ações de *marketing* mais eficientes e eficazes. Realizar a segmentação dos clientes a partir do valor futuro estimado que cada cliente irá gerar para a empresa é certamente uma alternativa bastante interessante para maximizar os retornos das ações de *marketing* realizadas. Sendo assim, os clientes foram ordenados a partir do valor do seu *CLV*, do cliente com maior valor ao cliente com menor valor e, após, foram subdivididos em decis, dando origem a 10 segmentos de clientes. Essa segmentação e o valor médio do *CLV* de cada segmento estão evidenciados no gráfico da Figura 18:

Figura 18 – Segmentação da base de clientes a partir o *CLV* - modelo aplicado



Diante dessa segmentação obtida, podemos verificar que, frente à amostra de clientes analisada, apenas 30% dos clientes provêm valor. Ao tomar conhecimento desse resultado, os profissionais da empresa estudada afirmaram que isso é coerente com a expectativa dos mesmos, uma vez que a experiência passada desses gestores indica que o percentual de clientes de toda que geram valor para a empresa está entre 20% e 30%.

A partir da segmentação obtida, o gestor de *marketing* passa a ter condições de planejar suas ações de *marketing* conhecendo cliente a cliente e tendo a estimativa de seu valor futuro. Esses 10 segmentos podem ainda ser reduzidos a um número menor de segmentos. Além disso, conhecendo o valor de cada cliente, os antigos segmentos utilizados pela empresa, como pessoa física ou pessoa jurídica, *business-to-business* ou *business-to-customer*, não precisam necessariamente ser deixados de lado. O gestor passa a ter ainda a alternativa de segmentar os clientes de cada um desses segmentos tradicionais a partir do seu valor futuro.

Com o modelo de *CLV* aplicado, para aquelas previsões consistentes, os clientes que desertaram e provavelmente não comprarão, os clientes novos que apresentam potencial de crescimento e os clientes antigos que compram frequentemente e que mantêm suas posições financeiras elevadas estarão alocados adequadamente em seus devidos segmentos. Assim, uma lista não esgotada de *insights* relacionados à gestão de clientes emerge:

- Os custos com catálogos, revistas entre outras ações de incentivo de *marketing* podem ser redirecionadas;

- Os clientes atendidos de forma personalizada por gestores de investimentos, que hoje são definidos apenas com base no tamanho da sua posição financeira, poderão também ser definidos em função dos segmentos estabelecidos;
- Os clientes dos segmentos com valor negativo podem ser direcionados para transacionar por canais de venda menos custosos e sem assessoramento personalizado;
- Aqueles clientes que reconhecidamente são bons clientes e que estão em segmentos com valor futuro baixo ou negativo podem estar deixando de comprar da empresa e migrando para um concorrente. Diante disso, ações para recuperar falhas que possam ter ocorrido ou simplesmente para reativar esses clientes podem ser tomadas;
- A partir da aplicação recorrente do modelo de *CLV*, a definição da remuneração variável dos agentes pode passar a levar em conta também o valor futuro de seus clientes e o segmento no qual eles estão;
- Os descontos concedidos aos clientes podem passar a ser definidos considerando também o segmento em que o cliente está;
- Utilizar os valores de *CLV* de cada cliente para traçar os perfis dos clientes com alto valor e dos clientes com baixo valor. Dessa forma, uma expectativa de valor futuro pode ser vislumbrada quando clientes novos são integrados a base de clientes.

Essas são apenas algumas das ações que podem ser tomadas pelo gestor de *marketing* uma vez que a previsão de rentabilidade de cada cliente passa a ser uma métrica utilizada. À medida que essa métrica é calculada recorrentemente e que os problemas de previsão vão sendo reconhecidos e solucionados, bem como as variáveis preditoras mais adequadas vão sendo consolidadas, o modelo aplicado em larga escala pode ser uma importante ferramenta de gestão que permite a definição de segmentos de clientes a partir da análise individual do valor futuro de cada cliente. Dessa forma, conforme requisitado por Ambler (2003), o conhecimento das fontes de receita da empresa e os *insights* a respeito de como aumentá-las passa a ser uma realidade. Além disso, o profissional de *marketing* passará a trabalhar com uma métrica *forward-looking* que traduz financeiramente os resultados obtidos pelas suas ações realizadas, criando, assim, uma alternativa para preencher a lacuna referente a dificuldade de comprovar produtividade das ações da área de *marketing*, principalmente aquelas relacionadas a gestão de clientes. Quando analisado de forma agregada, por meio do

valor do *CE*, as estimações realizadas pelo modelo possuem ainda o potencial para ser um argumento defendido junto aos acionistas para comprovar a geração de riqueza a partir das ações de *marketing* na medida em que, ao longo do tempo, o valor do *CE* for elevado.

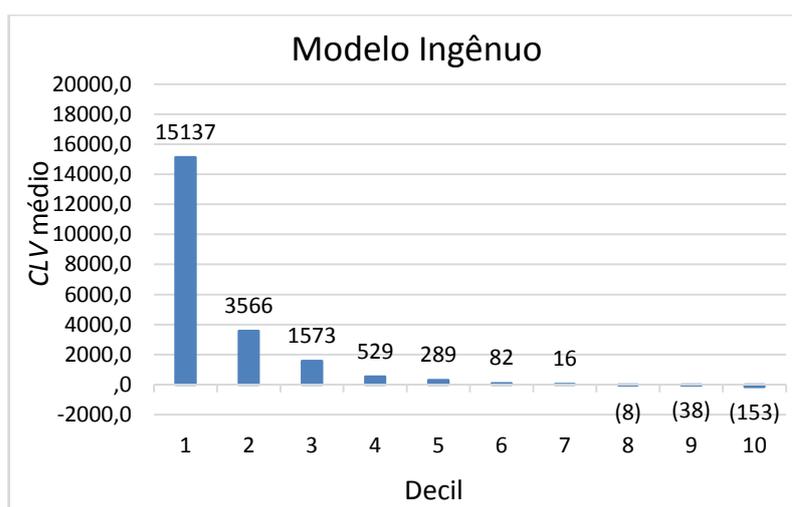
4.5 VERIFICAÇÃO DA CAPACIDADE DE PREVISÃO DO MODELO APLICADO

A cada possível modelo testado, 3 testes eram realizados para verificar a sua capacidade de previsão. A seguir, as verificações realizadas a partir desses testes serão apresentadas diante dos resultados do modelo final aplicado.

4.5.1 Comparação com o modelo ingênuo

Uma vez calculados os *CLVs* por meio do modelo denominado ingênuo, que conta apenas com as médias históricas dos custos de *marketing* e das margens de contribuição para definir os 36 fluxos de caixa futuros, o mesmo gráfico de segmentação da base de clientes feito para os resultados do modelo aplicado no presente trabalho também foi feito para os resultados do modelo ingênuo. Esse gráfico está apresentado a seguir na Figura 19:

Figura 19 – Segmentação da base de clientes a partir o *CLV* - modelo ingênuo



Embora diferenças possam ser verificadas, como o valor médio do *CLV* no modelo ingênuo que se aproxima do zero apenas no sétimo decil, enquanto no modelo aplicado esse valor já é negativo no quarto decil, quando os gráficos de segmentação da base clientes dos dois modelos concorrentes são analisados, o que mais chama a atenção é que o

comportamento dos *CLVs* dos clientes nos dois modelos é bastante semelhante. Não por coincidência, o comportamento dos *CLVs* dos clientes de duas empresas distintas analisadas por Kumar e Shah (2009) também foi bastante semelhante a esse dos dois modelos tratados no presente estudo. Em verdade, esse é um comportamento, salvo exceções, que é esperado, pois é comum ter-se uma parcela reduzida de clientes que trazem valor para a empresa. Além disso, o valor do *CE* dos clientes retidos é bastante semelhante nos dois modelos. Enquanto o CE_R do modelo aplicado, conforme já citado, é R\$ 205.463,20, o valor quando calculado pelo modelo ingênuo é R\$ 222.502,20.

Sendo assim, alguém pode perguntar: então por que empreender esforços para mensurar o *CLV*? A resposta é esclarecedora e, felizmente, animadora. Devemos empreender esforços para mensurar o *CLV*, principalmente sob o modo *bottom-up*, porque o comportamento dos gráficos apresentados evidencia apenas como o valor futuro está distribuído entre os clientes, mas não indica nada sobre quais clientes são esses. Diante desses dois gráficos, a próxima pergunta deveria ser: Quais clientes estão compondo cada um dos decis? Ao responder a essa pergunta, analisando a tabela presente no Anexo C, constata-se que os clientes que no passado agregaram grande valor, conforme é captado pelo modelo ingênuo, não serão necessariamente os clientes que futuramente agregarão maior valor, conforme é captado pelo modelo de *CLV* aplicado no presente estudo. O modelo ingênuo leva em consideração os valores médios gerados e gastos de cada cliente durante todo o período passado. Esse modelo não é sensível ao fato de algum cliente ter reduzido ou zerado sua posição financeira (*Net*), ter deixado de comprar há muitos meses, entre outras possibilidades. Diante disso, quando se utiliza um modelo que busca captar essas nuances, é possível identificar com maior precisão qual será o valor futuro de cada cliente.

Embora o Anexo D, proveniente da verificação da capacidade de previsão do modelo aplicado, apresente uma análise de cada um dos 100 clientes presentes na amostra, para exemplificar o motivo dessas diferenças nas previsões dos *CLVs*, é válido comparar os resultados obtidos pelos dois modelos para alguns dos clientes. O cliente de código 49, por exemplo, é um cliente que apresenta o 4º maior *CLV* quando analisado pelo modelo aplicado e é o 22º maior *CLV* quando analisado pelo modelo ingênuo. Essa é uma previsão consistente do modelo aplicado e é capaz de traduzir o potencial desse cliente, que realizou duas compras pequenas nos dois primeiros anos e, nos últimos seis meses observados, comprou em todos e apresentou uma margem de contribuição crescente ao longo desses meses, passando de R\$ 37,00 para R\$ 2547,00. A posição financeira desse cliente também cresceu bastante nesses

últimos seis meses passando de R\$ 19.000,00 para cerca de R\$ 400.000,00. Quando analisamos apenas o valor médio pelo modelo ingênuo, essa tendência de crescimento não é captada.

Ao analisar, em contrapartida, o cliente de código 28, diante dos resultados do modelo ingênuo, estaríamos colocando-o no 4º lugar, entre os clientes com maior *CLV*, no entanto, ao estimar-se o seu valor futuro por meio do modelo aplicado, constata-se que cliente possui *CLV* negativo. Essa diferença nos resultados é perfeitamente plausível quando, analisando-se as transações observadas desse cliente, constata-se que ele era um cliente realmente ótimo, pois realizou compras em 17 dos 29 períodos e possuía margens altas em torno de R\$ 1.000,00. No entanto sua última compra foi em 10/2012. Após, não comprou mais e sua posição financeira caiu de R\$ 282.000,00 no início de 2011 para R\$ 389,00 em 05/2013. A possibilidade de esse cliente ter desertado, se tornado inativo, é bem alta. Diante disso, a previsão de seu valor não pode continuar sendo aquela calcada em uma média que foi elevada há quase um ano. Frente a uma previsão ruim para o *CLV* desse cliente, os gestores de *marketing* poderão decidir se há possibilidade de tentar recuperar esse cliente ou se a solução é considerá-lo perdido e realmente entre aqueles com *CLV* negativo.

Diante disso, salvo os desvios de previsão encontrados e evidenciados na subseção seguinte, o modelo aplicado foi capaz de prever adequadamente o valor futuro dos clientes, sobrepondo-se aos resultados obtidos pelo modelo ingênuo. Além disso, comparando os dois modelos, foi possível constatar o ganho gerencial proporcionado pelo modelo aplicado, uma vez que é capaz de identificar as oscilações do comportamento de compra dos clientes ao longo do tempo.

4.5.2 Análise cliente a cliente

A análise mais profunda a respeito da capacidade de previsão do modelo foi realizada de forma analítica, cliente a cliente. Para tanto, analisaram-se os valores de *CLV* previstos para cada cliente frente ao seu histórico de transações para verificar a consistência das previsões realizadas. Nessa tarefa, foi possível identificar muitas fragilidades ainda não solucionadas do modelo final, no entanto foi possível constatar que, para aqueles casos nos quais as previsões se mostraram consistentes, o potencial gerencial do modelo é grandioso. Sendo assim, no Anexo D apresenta-se uma breve análise de cada cliente evidenciando-se os erros e os acertos de previsão ocorridos.

Diante dessas análises, que foram naturalmente resumidas em pequenos pareceres sobre cada cliente, podemos concluir que 84% por cento das previsões apresentam consistência. Ou seja, aqueles clientes que realmente parecem ter deixado de comprar, desertaram, foram identificados e seus *CLVs* foram negativos. Da mesma forma, clientes que tiveram poucas compras durante todo o período observado e que não são clientes novos, que entraram na base ao longo do período observado, também tiveram *CLV* negativo, pois nenhuma compra foi prevista. Por outro lado, os clientes com alto volume de compras e com algum potencial de geração de margem tiveram *CLV* positivo e os clientes novos, que tiveram poucas compras, mas todas nos últimos períodos e com tendência de crescimento do valor futuro, tiveram *CLV* positivo.

Enquanto os acertos são bastante animadores e são a maioria, o que indica uma boa capacidade de previsão do modelo, é preciso destacar também os pontos negativos que precisam ser aperfeiçoados. Os custos de *marketing*, conforme já mencionado na seção sobre a estimação dos parâmetros, gerou dificuldades para que a sua previsão fosse realizada de forma aceitável, porque seus dados foram apropriados *a posteriori* a partir da incidência de compra e do *net* observados de cada cliente. Ao testar-se a incidência de compra dos períodos anteriores (*Buy lag 1* e *Buy lag 2*) como variáveis preditoras, seu coeficiente ficou muito acima de 5, o que solucionou o problema de previsão dos clientes que desertaram, pois esses clientes passaram a ter custos futuros bem próximos de 0 ou até mesmo 0, no entanto distorceu a previsão dos clientes que tiveram compras no último período observado, pois, diante de um coeficiente alto, esses clientes tiveram seus custos de *marketing* elevados, acima de R\$ 100,00 ao mês, o que estaria bastante distante do valor máximo de custo de *marketing* de R\$ 28,15 praticado pela empresa no período observado. Diante disso, para buscar distorcer o menos possível o valor futuro dos clientes, optou-se pelo modelo que trouxe previsão dentro da faixa de valor aplicada no período observado. Apesar dessa solução adotada, outro problema ainda foi verificado nas previsões realizadas: existiam valores para os custos de *marketing* também para os clientes que não tinha incidência de compra prevista. Esse, portanto, é o motivo pelo qual existe uma quantidade elevada de clientes com *CLV* negativo, enquanto muitos deles poderiam ter *CLV* próximo de zero. Entende-se que os valores negativos previstos para os *CLVs* não foram extremos, por isso decidiu-se adotar esse modelo por apresentar maior parcimônia.

Outro ponto negativo observado nas previsões realizadas e destacado em algumas das análises dos clientes foi o fato de algumas margens de contribuição previstas terem sido

infladas, quando comparado com a margem de contribuição média do cliente no período observado. Apesar de esse ser evidentemente um problema que também precisa ser aperfeiçoado, entende-se que esses valores também não são extremos e, uma vez que o principal objetivo do modelo foi atingido na maioria dos casos, que é capturar o potencial do cliente, decidiu-se tolerar os casos em que as margens previstas foram elevadas.

4.5.3 Análise da previsão do período subsequente

O último teste de previsão realizado foi a verificação da previsão para o período $t + 1$. Para tanto, estimou-se novamente o modelo retirando-se o último período observado da base de dados. Dessa forma, os valores previstos para o período $t + 1$ foram comparados com os valores reais desse período.

Os resultados são positivos principalmente em relação à previsão da incidência de compra. O percentual de acerto foi de 80%, o que está bem próximo dos 84% de acerto responsáveis pela consistência das previsões para os *CLVs*.

Quando analisados os valores previstos para os custos de *marketing*, observou-se que todos os custos ficaram dentro do custo máximo de R\$ 16,73 praticado no período, o que demonstra que não houve elevação dos valores dos custos de *marketing* previstos. Também de forma positiva, 78% dos clientes 32 clientes que de fato realizaram compras receberam os custos de *marketing*, o que realmente deveria ocorrer diante da política de relacionamento adotada pela empresa. No entanto, evidenciando a dificuldade já mencionada de prever custos de *marketing* zero para os clientes que não realizarão mais compras, 55% dos 68 clientes que de fato não realizaram compras receberam algum valor de custos de *marketing*. Por fim, o total de custos de *marketing* previsto foi R\$ 404,52, enquanto o total praticado foi R\$ 524,67; a média dos custos de *marketing* previstos foi de R\$ 4,05, enquanto a média dos custos de *marketing* praticados foi de R\$ 5,25. Isso leva à conclusão de que a decisão pela parcimônia do modelo foi alcançada, uma vez que os custos de *marketing* extremos previstos por modelos alternativos não existiram.

Quando analisadas as previsões para as margens de contribuição, foi possível constatar que considerando o fato de haver ou não haver margem de contribuição praticada e seguindo o percentual de acerto para a incidência de compra, a previsão foi acertada em 81% desses casos. Os valores, porém, conforme já comentado sobre a necessidade de maior precisão nesse ponto, foram inflados, sendo que, somando o aumento marginal de cada cliente, o valor

total previsto foi R\$ 10.467,24 e o valor praticado foi R\$ 1.458,28. O valor médio previsto para a margem de contribuição foi de R\$ 128,13 e o valor médio praticado para a margem de contribuição foi de R\$ 14,58. Esse é um resultado negativo, pois demonstra que muitas margens de contribuição estão sendo estimadas com valores acima do que foi praticado.

5 CONCLUSÃO

5.1 CONSIDERAÇÕES FINAIS

Diante das dificuldades encontradas pelos profissionais de *marketing* para mensurar a produtividade de suas ações e relacionar esses resultados com a geração de valor para o acionista, a participação estratégica dessa área dentro das empresas foi reduzida. Diante desse cenário, esses mesmos profissionais, bem como os acadêmicos da área, buscam métricas mais adequadas para gerenciar os resultados de suas mais diversas ações. No que tange à gestão de clientes, especificamente, entende-se que, assim como o *marketing* tem a responsabilidade pela gestão do relacionamento com os clientes, deveria assumir também sua parcela de responsabilidade na gestão da lucratividade desses clientes. Isto requer, porém, que os profissionais da área não somente mensurem quanto os clientes geraram de lucratividade, mas o quanto se estima que cada um trará de rentabilidade no futuro.

Diante dessa lacuna existente na prática de *marketing* dentro das empresas, o *CLV* é apresentado como uma forma de contribuir substancialmente para a melhora desse panorama, uma vez que por meio do *CLV* é possível estimar o valor futuro dos clientes. O presente estudo e estudos anteriores comprovam que essa métrica, quando mensurada de forma consistente e a partir do modo *bottom-up*, é uma poderosa ferramenta capaz de traduzir a expectativa de rentabilidade de cada cliente, levando em consideração o seu comportamento de compra e os incentivos despendidos pela empresa para que cada cliente venha a comprar. Ao conhecer-se esse valor, os clientes podem ser segmentados de acordo com o seu valor futuro estimado e, diante disso, a alocação dos recursos disponíveis de *marketing* passa a ser mais precisa. A definição de perfis de clientes de acordo a sua perspectiva de rentabilidade, bastante utilizada por empresas, principalmente do setor financeiro, para segmentar os clientes novos sem histórico de compras, também pode ser incrementada quando os perfis são construídos a partir de uma métrica *forward-looking*. Além disso, ao considerar o *CE* como uma das métricas de seu *dashboard* que devem ser maximizadas, o gestor de *marketing* passa

a ter uma importante forma de evidenciar financeiramente os resultados de suas ações perante os acionistas.

Dada essa realidade animadora, uma questão parece ser chave para que todo esse potencial possa ser colocado em prática: como mensurar de forma adequada o *CLV*? É diante dessa questão que o presente trabalho foi conduzido com o objetivo de verificar os resultados da aplicação de um modelo estatisticamente sofisticado para mensurar o *CLV*. A pesquisa realizada enfrentou importantes desafios, como a compreensão de modelos estimados por meio de estatística bayesiana, que ainda carece de maior difusão na área de *marketing*, principalmente no território nacional, e a dificuldade de compreensão do método aplicado nos trabalhos de Kumar *et al.* (2008) e Kumar e Shah (2009), que foram, até certo ponto, lacônicos na apresentação das condicionais completas necessárias para estimar os parâmetros desses modelos. No entanto, acreditando-se no potencial da metodologia e do modelo construído por esses autores, foi possível compreender grande parte das lacunas deixadas e construir e aplicar um modelo baseado nesses trabalhos, bem como no trabalho de Cowles, Carlin e Connet (1996), que evidenciaram detalhadamente suas condicionais completas.

Os resultados obtidos evidenciaram que, nos casos em que o *CLV* foi mensurado de forma consistente, todo o seu potencial defendido na literatura pôde ser comprovado, pois a capacidade de identificar as oscilações no comportamento passado de compra dos clientes foi bastante satisfatória, e o valor futuro dos clientes refletiu de forma consistente inclusive as expectativas que podem ser estabelecidas para esses clientes quando são investigados caso a caso, de forma analítica.

Os resultados evidenciaram, porém, algumas limitações ainda não solucionadas do modelo que causaram inconsistências na previsão de 20% dos clientes analisados e distorções nos valores de algumas margens de contribuição previstas, o que acabou elevando o valor do *CLV* de alguns clientes. Além disso, o tempo de processamento do modelo por meio do código produzido está bastante aquém do ideal, e as soluções já vislumbradas precisam ser implementadas. Entende-se que essas e outras limitações que serão destacadas a seguir são parte de um processo de desenvolvimento científico. Diante da complexidade do modelo construído, acredita-se que apenas a continuidade do trabalho, buscando aperfeiçoar a ferramenta disponível será capaz de superar as limitações e potencializar os resultados bastante positivos que já puderam ser observados.

Por fim, o trabalho realizado revelou também um pré-requisito importante para a implementação bem-sucedida da estimação do *CLV*. Diante da pretensão do modo *bottom-up*

de estimar-se o *CLV* de cada cliente, a formação de uma base de dados que guarde as informações tanto transacionais quanto demográficas dos clientes é imprescindível. Essas informações incluem a margem de contribuição calculada a cada transação de forma automatizada, todos os contatos de *marketing* realizados junto ao cliente, bem como as variáveis que possam ser preditoras das três variáveis determinantes do *CLV* diante do conhecimento sobre o negócio de seus gestores e de análises prévias realizadas para identificar as variáveis com maior correlação com as variáveis determinantes do *CLV*.

5.2 LIMITAÇÕES DA PESQUISA E SUGESTÕES PARA PESQUISAS FUTURAS

Apesar do estágio alcançado, a pesquisa ainda possui limitações que podem e devem ser sanadas em pesquisas futuras. Os desvios observados nas previsões realizadas para os custos de *marketing*, embora causados também pelo fato de esses custos terem sido alocados *a posteriori*, e os valores superestimados das margens de contribuição são exemplos marcantes da fase ainda inicial de refinamento do modelo. Além disso, o tempo de processamento do modelo, conforme já mencionado, limita a aplicação em larga escala da estimação do *CLV* e, conseqüentemente, do valor integral do *CE* da organização estudada. Devido a esse problema, foi preciso utilizar uma amostra bastante reduzida para permitir o elevado número de testes necessários para o ajuste do modelo. Os esforços de programação já indicados ou outros que venham a ser vislumbrados devem ser empregados para que essa barreira seja superada.

Em relação aos hiperparâmetros utilizados, os interceptos, que são individuais por cliente, tiveram sua variância estimada, no entanto é possível estimar também a média desses interceptos, o que possibilitaria, em um nível hierárquico anterior, ponderar a média dos interceptos por meio das variáveis demográficas de cada cliente. Quanto à estimação dos coeficientes para as variáveis independentes, seria importante que eles também fossem individuais para cada cliente, com sua média e variância estimadas. Essas melhorias, já implementadas por Kumar e Shah (2009), provavelmente trariam maior capacidade de previsão ao modelo.

Outra limitação da pesquisa realizada, ainda em relação a estimação dos parâmetros, é que o vetor latente *Buy** é constante. Conforme Cowles, Carlin e Connet (1996), o vetor latente *Buy** poderia ser incluído como o primeiro passo do amostrador de Gibbs. Essa alternativa foi testada, no entanto não foi possível alcançar a convergência dos parâmetros da regressão referente ao vetor latente. Gelman *et al.* (2004) e Cowles, Carlin e Connet (1996)

indicam que fixar a variância da variável latente em 1 seria uma alternativa para solucionar o problema de identificabilidade existente. Essa alternativa realmente pode ser o caminho, entretanto os testes realizados também não apresentaram convergência dos parâmetros da regressão referente ao vetor latente. É importante que esforços futuros venham a solucionar essa limitação.

A base de dados obtida e utilizada possui apenas 29 períodos, o que certamente torna menos eficiente a tarefa de realizar a previsão de uma série de 36 períodos futuros. Além disso, algumas variáveis que seriam potenciais variáveis preditoras, como o resultado (ganho ou perda de capital) obtido pelos clientes ao longo do tempo, não puderam ser obtidas com a confiabilidade necessária e, portanto, foram excluídas da base de dados final.

Por fim, quanto à forma de estimação do *CE*, além de não representar o real *CE* de toda a base de clientes, ainda cabe crítica a sua fórmula de cálculo. Os valores futuros são estimados com base em uma estimativa arbitrária de crescimento da base, que deve ser definida pelos diretores da empresa. Pesquisas futuras poderiam também levar em consideração a estimação da taxa de aquisição futura a partir do histórico do que a empresa tem sido capaz de adquirir de clientes.

REFERÊNCIAS

AAKER, D. A. **Building strong brands**. Nova Iorque: Free Press, 1996a.

AAKER, D. A. Measuring brand equity across products and markets. **California Management Review**, v. 38, n. 3, p. 102-120, 1996b.

ALLENBY, G. **The Bayesian Revolution in Marketing Research**. Disponível em: <<http://www.msi.org/video/the-bayesian-revolution-in-marketing-research/>>. Acesso em: 03 jun. 2014.

AMBLER, T. **Marketing and the Bottom Line: The Marketing Metrics to Pump Up Cash Flow**. London: FT Prentice Hall, 2003.

ANDERSON, P. The *Marketing Management/Finance Interface*. **Proceedings of the 1979 Marketing Educator's Conference**. AMA Chicago, p. 325-329, 1979.

ANDERSON, E. W.; FORNELL, C.; MAZVANCHERYL, S. K. Customer Satisfaction and Shareholder Value. **Journal of Marketing**, v. 68, p. 172-185, 2004.

BAKER, S.; HOLT, S. Making marketers accountable: a failure of *Marketing* education? **Marketing Intelligence & Planning**, v. 22, n. 5, p. 557-567, 2004.

BERGER, P. D.; NASR, N. Customer Lifetime Value: *Marketing* models and applications. **Journal Interactive Marketing**, v. 12, n. 1, p. 17-30, 1998.

BERGER, P. D.; BOLTON, R. N.; BOWMAN, D.; BRIGGS, E.; KUMAR, V.; PARASURAMAN, A.; TERRY, C. *Marketing* Actions and the Value of Customer Assets: A Framework for Customer Asset Management. **Journal of Service Research**, v. 5, n. 1, p. 39-54, 2002.

BLATTBERG, R. C.; DEIGHTON, J. Manage *Marketing* by the customer equity test. **Harvard Business Review**, v. 74, n. 4, p. 136-144, 1996.

BLATTBERG, R. C.; MALTHOUSE, E. C.; NESLIN, S. A. Customer Lifetime Value: Empirical Generalizations and Some Conceptual Questions. **Journal of Interactive Marketing**, v. 23, p. 157-168, 2009.

BORGES, L. C. **Análise bayesiana do modelo fatorial dinâmico para um vetor de séries temporais usando distribuições elípticas**. Tese (Doutorado em Estatística) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2008.

BORLE, S.; SINGH, S. S.; JAIN, D. C. Customer Lifetime Value Measurement. **Management Science**, v. 54, n. 1, p. 100-112, 2008.

BOX, G. E. P. Science and Statistics. **Journal of the American Statistical Association**, v. 71, n. 356. p. 791-799, 1976.

BOX, G. E. P.; DRAPER, N. R. **Response surfaces, mixtures and ridge analyses**. Nova Jersey: John Wiley & Sons, 2007.

CAMPOS, P. S. S. **Estimação Bayesiana em modelos de estimação logística**. Dissertação (Mestrado em Matemática e Estatística) – Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém 2007.

COWLES, M. K.; CARLIN, B. P. Markov Chain Monte Carlo convergence diagnostics: a comparative review. **Journal of the American Statistics Association**, v. 91, n. 434, p. 883-904, 1996.

COWLES, M. K.; CARLIN, B. P.; CONNETT, J. E. Bayesian Tobit Modeling of Longitudinal Ordinal Clinical Trial Compliance Data with Nonignorable Missingsness. **Journal of the American Statistics Association**, v. 91, n. 433, p. 86-98, 1996.

DAY, G. *Marketing's* contribution to the strategy dialogue. **Journal of the Academy of Marketing Science**, v. 20, n. 4, 323-330, 1992.

DAY, G.; FAHEY, L. Valuing Market Strategies. **Journal of Marketing**, v. 52, p. 45-57, 1988.

DAY, G. S.; MONTGOMERY, D. B. Charting new directions for *Marketing*. **Journal of Marketing**, v. 63, p. 3-13, 1999.

DONKERS, B.; VERHOEF, P. C.; JONG, M. G. Modeling *CLV*: A test of competing models in the insurance industry. **Quantitative Marketing and Economics**, v. 5, p. 163-190, 2007.

DOYLE, P. Valuing *Marketing's* contribution. **European Management Journal**, v. 18, n. 3, p. 233–245, 2000.

FADER, P. S.; HARDIE, B. G. S.; LEE, K. L. “Counting Your Customers” the Easy Way: An Alternative to the Pareto/NBD Model. **Marketing Science**, v. 24, n. 2, p. 275-284, 2005.

FORNELL, C.; MITHAS, S.; MORGESON, F.; KRISHNAN, M. S. Customer Satisfaction and Stock Prices: High Returns, Low Risk. **Journal of Marketing**, v. 70, p. 3–14, 2006.

GELFAND, A. E.; SMITH, A. F. M. Sampling-Based Approaches to Calculating Marginal Densities. **Journal of the American Statistical Association**, v. 85, n. 410, p. 398-409, 1990.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian Data Analysis**. 2a ed. Nova Iorque: Chapman & Hall, 2004.

GEMAN, S.; GEMAN, D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. PAMI-6, n. 6, p. 721-741, 1984.

GEWEKE, J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In: Bernardo, J. M.; Berger, J.; David, A. P.; Smith, A. F. M. (org.). **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992.

GÖK, O.; HACIOGLU, G. The organizational roles of *Marketing* and *Marketing* managers. **Marketing Intelligence & Planning**, v. 28, n. 3, p. 291-309, 2010.

GUJARATI, D. N. **Econometria básica**. 3a ed. São Paulo: Makron Books, 2000.

GUPTA, S.; LEHMANN, D. R.; STUART, J. A. Valuing customers. **Journal of Marketing Research**, v. 41, p. 7–18, 2004.

GUPTA, S.; LEHMANN, D. R. Customer Lifetime Value and Firm Valuation. **Journal of Relationship Marketing**, v. 5, n. 2/3, p. 87-110, 2006.

HEIDELBERGER, P.; WELCH, P. D. Simulation Run Length Control in the Presence of an Initial Transient. **Operations Research**, v. 31, p. 1109-1144, 1983.

HOGAN, J. E.; LEMON, K. N.; RUST, R. R. Customer equity management: Charting new directions for the future of *Marketing*. **Journal of Service Research**, v. 5, n. 1, p. 4–12, 2002.

HOMBURG, C.; WORKMAN, J. P. Jr.; KROHMER, H. *Marketing's influence within the firm. Journal of Marketing*, v. 63, n. 2, p. 1-17, 1999.

JACKSON, D. Determining a customer's lifetime value. *Direct Marketing*, v. 51, n. 11, p. 60-62, 1989a.

JACKSON, D. Determining a customer's lifetime value (Part 2). *Direct Marketing*, v. 52, n. 1, p. 24-32, 1989b.

JACKSON, D. Insurance *Marketing*: Determining a customer's lifetime value (Part 3). *Direct Marketing*, v. 52, n. 4, p. 28-30, 1989c.

JAIN, D.; SINGH, S. S. Customer Lifetime Value Research in *Marketing*: a review and future directions. *Journal of Interactive Marketing*, v. 16, n. 2, p. 34-46, 2002.

JOSHI, A. M.; HANSENS, D. M. Movie Advertising and the Stock Market Valuation of studios. *Marketing Science*, v. 28, n. 2, p. 239-250, 2009.

KELLER, K. L. Conceptualizing, Measuring and managing customer-based brand equity. *Journal of Marketing*, v. 57, n. 1, p. 1-22, 1993.

KELLER, K. L. **Strategic Brand Management: building, measuring, and managing brand equity**. New Jersey: Prentice Hall, 1998

KUMAR, N. **Marketing como estratégia: uma orientação inovadora e comprovada para o crescimento e a inovação**. Rio de Janeiro: Elsevier, 2004.

KUMAR, V. **Managing Customers for Profit: Strategies to increase profits and build loyalty**. Upper Saddle River: Pearson Education/Wharton School Publishing, 2008.

KUMAR, V.; VENKATESAN, R.; BOHLING, T.; BECKMANN, D. The Power of *CLV*: Managing Customer Value at IBM. *Marketing Science*, v. 27, n. 4, p. 585-599, 2008.

KUMAR, V.; SHAH, D. Expanding the Role of *Marketing*: From Customer Equity to Market Capitalization. **Journal of Marketing**, v. 73, n. 6, p. 119-136, 2009.

LEHMANN, D. R. Metrics for making *Marketing* matter. **Journal of Marketing**, v. 68, n. 4, p. 73-75, 2004.

LEONE, R. P.; RAO, V. R.; KELLER, K. L.; LUO, A. M.; MCALISTER, L.; SRIVASTA, R. Linking brand equity to customer equity. **Journal of Service Research**, v. 9, n. 2, p. 125-138, 2006.

LEVITT, T. *Marketing* myopia. **Harvard Business Review**, v. 38, n. 4, p. 24-47, 1960.

MALTHOUSE, E. C.; BLATTBERG, R. C. Can We Predict Customer Lifetime Value? **Journal of Interactive Marketing**, v. 19, n. 1, p. 2-16, 2005.

MARKETING SCIENCE INSTITUTE (MSI). **Research Priorities 2008-2010**. Disponível em: <http://www.msi.org/pdf/MSI_RP08-10.pdf>. Acesso em: 15 jan. 2011.

MARKETING SCIENCE INSTITUTE (MSI). **Research Priorities 2014-2016**. Disponível em: <<http://www.msi.org/research/2014-2016-research-priorities/>>. Acesso em: 01 jun. 2014.

MIZIK, N.; JACOBSON, J. The Financial Value Impact of Perceptual Brand Attributes. **Journal of Marketing Research**, v. 45, p. 15-32, 2008.

MORGAN, N. A.; CLARK, B. H.; GOONER, R. Marketing productivity, Marketing audits, and systems for Marketing performance assessment: integrating multiple perspectives. **Journal of Business Research**, v. 55, n. 5, p. 363-375, 2002.

MSI WORKSHOP/CONFERENCE ON MARKETING METRICS. Swissôtel Washington The Watergate. **Five Parallel Small Group Working Sessions**, Washington, 1999.

PAULINO, C. D.; TURKMAN, M. A.; MURTEIRA, B. *Estatística Bayesiana*. Lisboa: Fundação Calouste Gulbenkian, 2003.

PAUWELS, K. H.; SILVA-RISSO, J. M.; SRINIVASAN, S.; HANSSSENS, D. M. New Products, Sales Promotions, and Firm Value: The Case of the Automobile Industry. **Journal of Marketing**, v. 68, p. 142-156, 2004.

PFEIFER, P. E.; CARRAWAY, R. L. Modeling Customer Relationships As Markov Chains. **Journal of Interactive Marketing**, v. 14, n. 2, p. 43-55, 2000.

RAO, V. R.; AGARWAL, M. K.; DAHLOFF, D. How Is Manifest Branding Strategy Related to the Intangible Value of a Corporation? **Journal of Marketing**, v. 68, p. 126-141, 2004

REINARTZ, W. J.; KUMAR, V. The Impact of Customer Relationship Characteristics on Profitable Life Duration. **Journal of Marketing**, v. 67, p. 77-99, 2003.

ROSSI, P. E.; ALLENBY, G. M. Bayesian Statistics and *Marketing*. **Marketing Science**, v. 22, n. 3, p. 304-328, 2003.

RUST, R. T.; LEMON, K. N.; ZEITHAML, V. A. Modeling Customer Equity. **Marketing Science Institute**. Working Paper Series, v. 01, n. 108, 2001.

RUST, R. T.; LEMON, K. N.; ZEITHAML, V. A. Return on *Marketing*: using customer equity to focus *Marketing* strategy. **Journal of Marketing**, v. 68, n. 1, p. 109-127, 2004.

RUST, R. T.; AMBLER, T.; CARPENTER, G. S.; KUMAR, V.; SRIVASTAVA, R. K. Measuring *Marketing* productivity: current knowledge and future directions. **Journal of Marketing**, v. 68, n. 4, p. 76-89, 2004

SCHMITTLEIN, D. C.; MORRISON, D. G.; COLOMBO, R. Counting your customers: Who are they and what will they do next? **Management Science**, v. 33, p. 1-24, 1987.

SCHMITTLEIN, D. C., PETERSON, R. A. Customer Base Analysis: An Industrial Purchase Process Application. *Marketing Science*, v. 13, n. 1, p. 41-68, 1994.

SHETH, J. N.; SISODIA, R. Feeling the Heat. *Marketing Management*, v. 4, p. 9-23, 1995a.

SHETH, J. N.; SISODIA, R. Feeling the Heat – Part 2. *Marketing Management*, v. 4, p. 19-33, 1995b.

SHETH, J. N.; SISODIA, R. High Performance *Marketing*. *Marketing Management*, v. 10, n. 3, p. 18-23, 2001.

SHETH, J. N.; SISODIA, R. *Marketing* productivity: issues and analysis. *Journal of Business Research*, v. 55, n. 5, p. 349-362, 2002.

SINGH, S. S.; JAIN, D. C. **Measuring Customer Lifetime Value: Models and Analysis**. Working Paper, INSEAD, Fontainebleau, 2010.

SORESCU, A.; SHANKAR, V.; KUSHWAHA, T. New Product Preannouncements and Shareholder Value: Don't Make Promises You Can't Keep. *Journal of Marketing Research*, v. 44, p. 468–489, 2007.

SRINIVASAN, S.; HANSSENS, D. M. *Marketing* and Firm Value: Metrics, Methods, Findings, and Future Directions. *Journal of Marketing Research*, v. 46, p. 293-312, 2009

SRIVASTAVA, R. K.; SHERVANI, T. A.; FAHEY, L. Market-based assets and shareholder value: a framework for analysis. *Journal of Marketing*, v. 62, p. 2-18, 1998.

STEWART, D. W. *Marketing* accountability: linking *Marketing* actions to financial results. *Journal of Business Research*, v. 62, n. 6, p. 636-643, 2009.

TANNER, T. A.; WONG, W. H. The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, v. 82, p. 528-550, 1987.

VENKATESAN, R.; KUMAR, V. A Customer Lifetime Value Framework for Customer Selection and Optimal Resource Allocation Strategy. **Journal of Marketing**, v. 68, p. 106-125, 2004.

VENKATESAN, R.; KUMAR, V.; BOHLING, T. Optimal Customer Relationship Management Using Bayesian Decision Theory: An Application for Customer Selection. **Journal of Marketing Research**, v. 44, p. 579-594, 2007.

VILLANUEVA, J.; HANSSENS, D. M. Customer Equity: Measurement, Management and Research Opportunities. **Foundations and Trends in Marketing**, v. 1, n. 1, p. 1-95, 2007.

WEBSTER, F.E. The rediscovery of the *Marketing* concept. **Business Horizons**, v. 31, n. 3, p. 29-39, 1988.

WEBSTER F. E. The changing role of *Marketing* in the corporation. **Journal of Marketing**, v. 56, n. 4, p. 1-17, 1992.

WEBSTER, F.E. Defining the new *Marketing* concept. **Marketing Management**, v. 2, n. 4, p. 22-31, 1994.

WEBSTER, F.E. *Marketing* management in changing times. **Marketing Management**, v. 11, n. 1, p. 18-23, 2002.

WEBSTER, F. E. A perspective on the evolution of *Marketing* management. **Journal of Public Policy & Marketing**, v. 24, n. 1, p. 121-126, 2005.

WEBSTER, F. E.; MALTER, A. J.; GANESAN, S. **Can Marketing regain its seat at the table?** Working Paper. Cambridge: *Marketing* Science Institute, Report n. 03, 2003.

WEBSTER, F. E.; MALTER, A. J.; GANESAN, S. The Decline and Dispersion of *Marketing* Competence. **MIT Sloan Management Review**, v. 46, n. 4, p. 35-43, 2005.

YOO, S.; HANSSENS, D. M. **Modeling the sales and customer equity effects of the *Marketing mix***. Working Paper, Universidade da Califórnia, Los Angeles, Anderson School of Management, 2005.

ZELLNER, A. **An introduction to Bayesian inference in econometrics**. New York: Wiley, 1971.

ANEXOS

ANEXO A – MÉDIAS E DESVIOS PADRÃO DOS INTERCEPTOS

Intercepto Regressão 1	Média	Desvio Padrão	Intercepto Regressão 2	Média	Desvio Padrão	Intercepto Regressão 3	Média	Desvio Padrão
Alfa 1.1	-0.198	0.158	Alfa 2.1	-0.379	0.030	Alfa 3.1	-0.124	0.005
Alfa 1.2	-0.169	0.154	Alfa 2.2	-0.307	0.030	Alfa 3.2	-0.151	0.007
Alfa 1.3	-0.167	0.160	Alfa 2.3	-0.417	0.032	Alfa 3.3	-0.159	0.003
Alfa 1.4	0.529	0.162	Alfa 2.4	-0.868	0.036	Alfa 3.4	-0.777	0.009
Alfa 1.5	0.079	0.150	Alfa 2.5	-0.366	0.031	Alfa 3.5	-0.172	0.015
Alfa 1.6	0.003	0.144	Alfa 2.6	-0.299	0.029	Alfa 3.6	-0.148	0.008
Alfa 1.7	0.004	0.149	Alfa 2.7	-0.462	0.032	Alfa 3.7	-0.146	0.016
Alfa 1.8	-0.190	0.164	Alfa 2.8	-0.171	0.035	Alfa 3.8	-0.131	0.010
Alfa 1.9	-0.137	0.159	Alfa 2.9	-0.569	0.032	Alfa 3.9	-0.347	0.063
Alfa 1.10	-0.053	0.155	Alfa 2.10	-0.298	0.030	Alfa 3.10	-0.147	0.011
Alfa 1.11	-0.206	0.156	Alfa 2.11	-0.300	0.030	Alfa 3.11	-0.131	0.018
Alfa 1.12	-0.079	0.150	Alfa 2.12	-0.454	0.037	Alfa 3.12	-0.193	0.023
Alfa 1.13	-0.194	0.166	Alfa 2.13	-0.464	0.036	Alfa 3.13	-0.202	0.006
Alfa 1.14	-0.161	0.158	Alfa 2.14	-0.519	0.032	Alfa 3.14	-0.159	0.004
Alfa 1.15	0.129	0.147	Alfa 2.15	-0.349	0.031	Alfa 3.15	-0.293	0.017
Alfa 1.16	0.386	0.160	Alfa 2.16	-0.213	0.043	Alfa 3.16	-0.172	0.010
Alfa 1.17	-0.134	0.152	Alfa 2.17	-0.461	0.031	Alfa 3.17	-0.324	0.004
Alfa 1.18	-0.086	0.150	Alfa 2.18	-0.526	0.031	Alfa 3.18	-0.437	0.032
Alfa 1.19	-0.209	0.154	Alfa 2.19	-0.476	0.030	Alfa 3.19	-0.136	0.002
Alfa 1.20	0.485	0.174	Alfa 2.20	-0.302	0.037	Alfa 3.20	-0.387	0.005
Alfa 1.21	-0.208	0.158	Alfa 2.21	-0.233	0.032	Alfa 3.21	-0.135	0.005
Alfa 1.22	-0.003	0.151	Alfa 2.22	-0.518	0.030	Alfa 3.22	-0.164	0.009

Alfa 1.23	-0.097	0.151	Alfa 2.23	-0.387	0.033	Alfa 3.23	-0.209	0.008
Alfa 1.24	-0.069	0.150	Alfa 2.24	-0.224	0.033	Alfa 3.24	-0.172	0.015
Alfa 1.25	0.433	0.160	Alfa 2.25	-0.154	0.032	Alfa 3.25	-0.224	0.013
Alfa 1.26	-0.209	0.152	Alfa 2.26	-0.139	0.044	Alfa 3.26	-0.121	0.005
Alfa 1.27	0.177	0.203	Alfa 2.27	-0.321	0.037	Alfa 3.27	-0.992	0.078
Alfa 1.28	0.034	0.156	Alfa 2.28	-0.325	0.038	Alfa 3.28	-0.263	0.009
Alfa 1.29	-0.197	0.153	Alfa 2.29	-0.258	0.030	Alfa 3.29	-0.123	0.003
Alfa 1.30	-0.077	0.155	Alfa 2.30	-0.466	0.030	Alfa 3.30	-0.456	0.006
Alfa 1.31	-0.121	0.155	Alfa 2.31	-0.235	0.031	Alfa 3.31	-0.198	0.013
Alfa 1.32	0.104	0.146	Alfa 2.32	-0.276	0.032	Alfa 3.32	-0.242	0.004
Alfa 1.33	-0.115	0.151	Alfa 2.33	-0.087	0.033	Alfa 3.33	-0.170	0.025
Alfa 1.34	-0.212	0.147	Alfa 2.34	-0.511	0.029	Alfa 3.34	-0.089	0.003
Alfa 1.35	-0.231	0.148	Alfa 2.35	-0.376	0.033	Alfa 3.35	-0.131	0.007
Alfa 1.36	0.016	0.153	Alfa 2.36	-0.281	0.030	Alfa 3.36	-0.500	0.017
Alfa 1.37	-0.166	0.148	Alfa 2.37	-0.330	0.029	Alfa 3.37	-0.154	0.006
Alfa 1.38	0.026	0.161	Alfa 2.38	-0.496	0.033	Alfa 3.38	-0.158	0.024
Alfa 1.39	-0.085	0.154	Alfa 2.39	-0.490	0.031	Alfa 3.39	-0.169	0.008
Alfa 1.40	-0.261	0.149	Alfa 2.40	-0.427	0.030	Alfa 3.40	0.365	0.008
Alfa 1.41	0.197	0.150	Alfa 2.41	-0.131	0.032	Alfa 3.41	-0.158	0.007
Alfa 1.42	0.076	0.152	Alfa 2.42	-0.352	0.036	Alfa 3.42	-0.204	0.013
Alfa 1.43	-0.102	0.152	Alfa 2.43	-0.325	0.030	Alfa 3.43	-0.222	0.002
Alfa 1.44	-0.124	0.151	Alfa 2.44	-0.506	0.030	Alfa 3.44	-0.125	0.008
Alfa 1.45	0.295	0.160	Alfa 2.45	-0.433	0.037	Alfa 3.45	-0.232	0.010
Alfa 1.46	-0.208	0.153	Alfa 2.46	-0.381	0.029	Alfa 3.46	-0.130	0.002
Alfa 1.47	-0.274	0.155	Alfa 2.47	-0.444	0.030	Alfa 3.47	-0.040	0.004
Alfa 1.48	-0.234	0.157	Alfa 2.48	-0.384	0.031	Alfa 3.48	-0.169	0.014
Alfa 1.49	-0.056	0.146	Alfa 2.49	-0.555	0.030	Alfa 3.49	-0.218	0.005
Alfa 1.50	-0.152	0.156	Alfa 2.50	-0.253	0.030	Alfa 3.50	-0.236	0.005

Alfa 1.51	0.153	0.156	Alfa 2.51	-0.304	0.032	Alfa 3.51	-0.163	0.007
Alfa 1.52	0.124	0.149	Alfa 2.52	-0.635	0.031	Alfa 3.52	-0.445	0.008
Alfa 1.53	0.163	0.152	Alfa 2.53	-0.449	0.038	Alfa 3.53	-0.528	0.037
Alfa 1.54	-0.207	0.153	Alfa 2.54	-0.490	0.030	Alfa 3.54	-0.147	0.006
Alfa 1.55	0.381	0.167	Alfa 2.55	-0.103	0.044	Alfa 3.55	-0.155	0.022
Alfa 1.56	0.322	0.150	Alfa 2.56	-0.329	0.032	Alfa 3.56	-0.393	0.029
Alfa 1.57	-0.188	0.155	Alfa 2.57	-0.224	0.030	Alfa 3.57	-0.140	0.002
Alfa 1.58	-0.179	0.149	Alfa 2.58	-0.563	0.029	Alfa 3.58	-0.131	0.007
Alfa 1.59	0.072	0.152	Alfa 2.59	-0.339	0.033	Alfa 3.59	-0.143	0.025
Alfa 1.60	-0.184	0.151	Alfa 2.60	-0.226	0.030	Alfa 3.60	-0.124	0.011
Alfa 1.61	0.513	0.173	Alfa 2.61	0.017	0.037	Alfa 3.61	-0.350	0.006
Alfa 1.62	0.410	0.157	Alfa 2.62	-0.041	0.032	Alfa 3.62	-0.270	0.021
Alfa 1.63	-0.258	0.149	Alfa 2.63	-0.510	0.030	Alfa 3.63	-0.061	0.003
Alfa 1.64	-0.232	0.159	Alfa 2.64	-0.503	0.032	Alfa 3.64	-0.149	0.008
Alfa 1.65	-0.013	0.150	Alfa 2.65	-0.309	0.035	Alfa 3.65	-0.248	0.007
Alfa 1.66	0.039	0.151	Alfa 2.66	-0.504	0.033	Alfa 3.66	-0.233	0.012
Alfa 1.67	0.142	0.161	Alfa 2.67	-0.309	0.036	Alfa 3.67	-0.214	0.007
Alfa 1.68	-0.238	0.154	Alfa 2.68	-0.585	0.031	Alfa 3.68	-0.131	0.010
Alfa 1.69	0.028	0.160	Alfa 2.69	-0.425	0.031	Alfa 3.69	-1.037	0.033
Alfa 1.70	0.092	0.149	Alfa 2.70	-0.297	0.036	Alfa 3.70	-0.147	0.034
Alfa 1.71	-0.121	0.149	Alfa 2.71	-0.479	0.031	Alfa 3.71	-0.161	0.004
Alfa 1.72	-0.195	0.149	Alfa 2.72	-0.167	0.030	Alfa 3.72	-0.126	0.002
Alfa 1.73	0.201	0.154	Alfa 2.73	-0.539	0.034	Alfa 3.73	-0.444	0.014
Alfa 1.74	-0.189	0.154	Alfa 2.74	-0.177	0.030	Alfa 3.74	-0.140	0.004
Alfa 1.75	-0.198	0.155	Alfa 2.75	-0.339	0.031	Alfa 3.75	-0.153	0.007
Alfa 1.76	0.045	0.147	Alfa 2.76	-0.406	0.037	Alfa 3.76	-0.305	0.017
Alfa 1.77	-0.135	0.142	Alfa 2.77	-0.317	0.030	Alfa 3.77	-0.129	0.003
Alfa 1.78	-0.134	0.145	Alfa 2.78	-0.227	0.029	Alfa 3.78	-0.179	0.015

Alfa 1.79	-0.252	0.157	Alfa 2.79	-0.346	0.030	Alfa 3.79	-0.130	0.014
Alfa 1.80	0.579	0.161	Alfa 2.80	-0.471	0.034	Alfa 3.80	-1.169	0.019
Alfa 1.81	0.494	0.163	Alfa 2.81	-0.329	0.035	Alfa 3.81	-0.536	0.011
Alfa 1.82	0.164	0.154	Alfa 2.82	-0.409	0.031	Alfa 3.82	-0.633	0.004
Alfa 1.83	0.356	0.161	Alfa 2.83	-0.446	0.032	Alfa 3.83	-0.235	0.017
Alfa 1.84	0.543	0.174	Alfa 2.84	0.044	0.044	Alfa 3.84	-0.160	0.010
Alfa 1.85	-0.082	0.156	Alfa 2.85	-0.456	0.033	Alfa 3.85	-0.450	0.033
Alfa 1.86	0.230	0.163	Alfa 2.86	-0.886	0.031	Alfa 3.86	-0.437	0.008
Alfa 1.87	0.406	0.159	Alfa 2.87	-0.766	0.032	Alfa 3.87	-1.072	0.013
Alfa 1.88	0.324	0.168	Alfa 2.88	-0.192	0.036	Alfa 3.88	-0.530	0.010
Alfa 1.89	-0.214	0.146	Alfa 2.89	-0.641	0.029	Alfa 3.89	-0.125	0.012
Alfa 1.90	-0.191	0.153	Alfa 2.90	-0.677	0.038	Alfa 3.90	-0.154	0.020
Alfa 1.91	0.309	0.154	Alfa 2.91	-0.252	0.035	Alfa 3.91	-0.680	0.026
Alfa 1.92	0.284	0.165	Alfa 2.92	-0.240	0.037	Alfa 3.92	-0.161	0.013
Alfa 1.93	-0.240	0.155	Alfa 2.93	-0.751	0.032	Alfa 3.93	0.114	0.018
Alfa 1.94	-0.021	0.151	Alfa 2.94	-0.338	0.031	Alfa 3.94	-0.194	0.005
Alfa 1.95	-0.157	0.152	Alfa 2.95	-0.183	0.030	Alfa 3.95	-0.132	0.013
Alfa 1.96	0.137	0.155	Alfa 2.96	-0.414	0.039	Alfa 3.96	-0.371	0.030
Alfa 1.97	-0.013	0.151	Alfa 2.97	-0.457	0.031	Alfa 3.97	-0.167	0.009
Alfa 1.98	-0.157	0.151	Alfa 2.98	-0.548	0.037	Alfa 3.98	-0.187	0.008
Alfa 1.99	0.166	0.159	Alfa 2.99	-0.261	0.034	Alfa 3.99	-0.152	0.011
Alfa 1.100	0.328	0.146	Alfa 2.100	0.028	0.036	Alfa 3.100	-0.265	0.006

ANEXO B – RESULTADOS DOS CRITÉRIOS DE CONVERGÊNCIAS PARA OS INTERCEPTOS

	Geweke	Heidelberger e Welch	<i>Halfwidth</i>		Geweke	Heidelberger e Welch	<i>Halfwidth</i>		Geweke	Heidelberger e Welch	<i>Halfwidth</i>
Intercepto	<i>Z-Score</i>	<i>p-value</i>	-	Intercepto	<i>Z-Score</i>	<i>p-value</i>	-	Intercepto	<i>Z-Score</i>	<i>p-value</i>	-
Alfa 1.1	-1.984	0.520	0.013	Alfa 2.1	0.039	0.687	0.002	Alfa 3.1	-1.168	0.403	0.000
Alfa 1.2	0.395	0.180	0.012	Alfa 2.2	2.034	0.208	0.002	Alfa 3.2	0.453	0.099	0.000
Alfa 1.3	0.658	0.087	0.011	Alfa 2.3	1.735	0.091	0.002	Alfa 3.3	0.861	0.053	0.000
Alfa 1.4	1.034	0.526	0.013	Alfa 2.4	1.241	0.913	0.003	Alfa 3.4	1.406	0.140	0.002
Alfa 1.5	0.482	0.696	0.012	Alfa 2.5	1.273	0.655	0.002	Alfa 3.5	-1.172	0.163	0.002
Alfa 1.6	0.476	0.083	0.012	Alfa 2.6	1.014	0.053	0.002	Alfa 3.6	-1.231	0.090	0.000
Alfa 1.7	0.907	0.057	0.013	Alfa 2.7	1.481	0.059	0.003	Alfa 3.7	-1.233	0.127	0.002
Alfa 1.8	-0.669	0.729	0.013	Alfa 2.8	1.610	0.462	0.003	Alfa 3.8	0.956	0.504	0.000
Alfa 1.9	1.642	0.157	0.016	Alfa 2.9	2.958	0.391	0.003	Alfa 3.9	-1.196	0.685	0.027
Alfa 1.10	-0.649	0.162	0.012	Alfa 2.10	0.883	0.381	0.003	Alfa 3.10	-1.071	0.065	0.002
Alfa 1.11	-1.045	0.310	0.015	Alfa 2.11	0.187	0.168	0.002	Alfa 3.11	-1.004	0.142	0.000
Alfa 1.12	-0.204	0.598	0.011	Alfa 2.12	1.560	0.577	0.002	Alfa 3.12	-1.077	0.292	0.004
Alfa 1.13	-0.828	0.436	0.013	Alfa 2.13	1.494	0.278	0.003	Alfa 3.13	0.805	0.797	0.000
Alfa 1.14	0.270	0.596	0.013	Alfa 2.14	1.399	0.214	0.003	Alfa 3.14	-1.518	0.373	0.000
Alfa 1.15	1.196	0.465	0.012	Alfa 2.15	1.478	0.673	0.002	Alfa 3.15	-1.394	0.130	0.002
Alfa 1.16	1.245	0.504	0.009	Alfa 2.16	1.002	0.544	0.002	Alfa 3.16	-1.203	0.487	0.000
Alfa 1.17	0.539	0.904	0.012	Alfa 2.17	1.722	0.967	0.002	Alfa 3.17	-1.546	0.395	0.000
Alfa 1.18	1.424	0.070	0.013	Alfa 2.18	2.038	0.154	0.002	Alfa 3.18	-1.176	0.474	0.010
Alfa 1.19	0.272	0.668	0.012	Alfa 2.19	1.767	0.051	0.003	Alfa 3.19	-2.177	0.693	0.000
Alfa 1.20	1.140	0.050	0.023	Alfa 2.20	1.087	0.557	0.004	Alfa 3.20	-1.026	0.496	0.000
Alfa 1.21	-0.334	0.502	0.014	Alfa 2.21	2.077	0.135	0.003	Alfa 3.21	-1.112	0.575	0.000
Alfa 1.22	-0.860	0.572	0.011	Alfa 2.22	0.603	0.233	0.003	Alfa 3.22	-0.246	0.482	0.000
Alfa 1.23	-0.189	0.917	0.012	Alfa 2.23	1.324	0.093	0.004	Alfa 3.23	-1.033	0.958	0.000

Alfa 1.24	0.304	0.121	0.012	Alfa 2.24	1.502	0.085	0.003	Alfa 3.24	-1.163	0.069	0.003
Alfa 1.25	0.667	0.193	0.020	Alfa 2.25	0.842	0.052	0.005	Alfa 3.25	-1.158	0.434	0.000
Alfa 1.26	-1.041	0.239	0.013	Alfa 2.26	1.132	0.474	0.002	Alfa 3.26	-1.021	0.491	0.000
Alfa 1.27	-1.177	0.490	0.014	Alfa 2.27	1.293	0.488	0.003	Alfa 3.27	1.042	0.659	0.023
Alfa 1.28	1.092	0.539	0.012	Alfa 2.28	1.488	0.326	0.002	Alfa 3.28	-1.397	0.472	0.001
Alfa 1.29	-1.660	0.097	0.014	Alfa 2.29	0.332	0.250	0.003	Alfa 3.29	0.676	0.065	0.000
Alfa 1.30	-0.299	0.659	0.012	Alfa 2.30	0.953	0.269	0.002	Alfa 3.30	-0.950	0.285	0.000
Alfa 1.31	0.467	0.828	0.012	Alfa 2.31	1.343	0.206	0.003	Alfa 3.31	1.112	0.707	0.000
Alfa 1.32	-0.809	0.467	0.012	Alfa 2.32	0.704	0.366	0.003	Alfa 3.32	-1.337	0.314	0.000
Alfa 1.33	-0.526	0.526	0.012	Alfa 2.33	1.134	0.581	0.002	Alfa 3.33	-1.365	0.132	0.004
Alfa 1.34	-0.616	0.306	0.012	Alfa 2.34	1.205	0.751	0.002	Alfa 3.34	-1.777	0.815	0.000
Alfa 1.35	-0.296	0.177	0.013	Alfa 2.35	1.442	0.100	0.003	Alfa 3.35	-1.131	0.094	0.000
Alfa 1.36	-2.388	0.524	0.013	Alfa 2.36	0.272	0.571	0.002	Alfa 3.36	0.920	0.342	0.005
Alfa 1.37	-0.901	0.223	0.013	Alfa 2.37	0.536	0.086	0.002	Alfa 3.37	-1.101	0.056	0.000
Alfa 1.38	-1.236	0.945	0.014	Alfa 2.38	0.436	0.291	0.003	Alfa 3.38	-1.123	0.190	0.003
Alfa 1.39	-0.077	0.878	0.012	Alfa 2.39	0.858	0.392	0.003	Alfa 3.39	-1.208	0.759	0.000
Alfa 1.40	0.979	0.154	0.012	Alfa 2.40	1.595	0.333	0.003	Alfa 3.40	-1.341	0.085	0.000
Alfa 1.41	1.418	0.153	0.013	Alfa 2.41	1.269	0.453	0.002	Alfa 3.41	-1.049	0.383	0.000
Alfa 1.42	-0.513	0.845	0.013	Alfa 2.42	0.822	0.931	0.002	Alfa 3.42	-0.885	0.066	0.002
Alfa 1.43	-0.923	0.183	0.012	Alfa 2.43	0.900	0.361	0.003	Alfa 3.43	-0.496	0.531	0.000
Alfa 1.44	-0.130	0.515	0.012	Alfa 2.44	1.621	0.491	0.002	Alfa 3.44	0.476	0.258	0.000
Alfa 1.45	0.656	0.443	0.015	Alfa 2.45	1.206	0.657	0.003	Alfa 3.45	-0.713	0.857	0.000
Alfa 1.46	0.054	0.741	0.012	Alfa 2.46	2.743	0.112	0.002	Alfa 3.46	-2.814	0.752	0.000
Alfa 1.47	-2.601	0.209	0.012	Alfa 2.47	-0.316	0.592	0.002	Alfa 3.47	1.728	0.365	0.000
Alfa 1.48	-2.292	0.129	0.013	Alfa 2.48	0.151	0.492	0.002	Alfa 3.48	-1.062	0.271	0.000
Alfa 1.49	0.282	0.395	0.012	Alfa 2.49	1.721	0.333	0.002	Alfa 3.49	0.823	0.466	0.000
Alfa 1.50	-2.222	0.084	0.013	Alfa 2.50	0.171	0.218	0.002	Alfa 3.50	1.225	0.085	0.000
Alfa 1.51	1.107	0.267	0.013	Alfa 2.51	1.019	0.611	0.003	Alfa 3.51	-0.767	0.669	0.000

Alfa 1.52	2.291	0.276	0.012	Alfa 2.52	1.650	0.709	0.002	Alfa 3.52	-1.046	0.346	0.000
Alfa 1.53	-0.220	0.785	0.009	Alfa 2.53	0.842	0.666	0.002	Alfa 3.53	-1.131	0.351	0.008
Alfa 1.54	-0.050	0.766	0.012	Alfa 2.54	2.404	0.361	0.002	Alfa 3.54	-1.117	0.420	0.000
Alfa 1.55	0.453	0.111	0.017	Alfa 2.55	0.773	0.081	0.003	Alfa 3.55	-1.148	0.213	0.004
Alfa 1.56	0.748	0.782	0.012	Alfa 2.56	0.939	0.721	0.002	Alfa 3.56	-1.080	0.343	0.005
Alfa 1.57	-1.710	0.075	0.012	Alfa 2.57	0.437	0.449	0.003	Alfa 3.57	-1.185	0.233	0.000
Alfa 1.58	-0.929	0.324	0.012	Alfa 2.58	1.266	0.127	0.002	Alfa 3.58	-1.228	0.216	0.000
Alfa 1.59	0.165	0.932	0.012	Alfa 2.59	0.815	0.064	0.004	Alfa 3.59	-0.954	0.262	0.004
Alfa 1.60	-1.373	0.630	0.012	Alfa 2.60	0.179	0.485	0.002	Alfa 3.60	-1.204	0.433	0.000
Alfa 1.61	0.787	0.073	0.016	Alfa 2.61	0.881	0.331	0.003	Alfa 3.61	-1.134	0.321	0.000
Alfa 1.62	0.629	0.279	0.013	Alfa 2.62	0.880	0.080	0.003	Alfa 3.62	-1.082	0.167	0.004
Alfa 1.63	1.046	0.194	0.012	Alfa 2.63	1.663	0.290	0.003	Alfa 3.63	-0.110	0.290	0.000
Alfa 1.64	-0.856	0.588	0.016	Alfa 2.64	0.992	0.455	0.003	Alfa 3.64	-1.220	0.745	0.000
Alfa 1.65	-0.995	0.438	0.012	Alfa 2.65	0.927	0.214	0.004	Alfa 3.65	-0.778	0.796	0.000
Alfa 1.66	-1.111	0.804	0.013	Alfa 2.66	0.564	0.684	0.003	Alfa 3.66	-1.195	0.173	0.002
Alfa 1.67	0.591	0.378	0.013	Alfa 2.67	1.008	0.231	0.003	Alfa 3.67	-0.977	0.085	0.000
Alfa 1.68	-0.562	0.240	0.012	Alfa 2.68	1.509	0.085	0.003	Alfa 3.68	-1.120	0.080	0.000
Alfa 1.69	-0.543	0.293	0.016	Alfa 2.69	0.770	0.103	0.003	Alfa 3.69	1.185	0.366	0.011
Alfa 1.70	-0.365	0.694	0.012	Alfa 2.70	0.981	0.061	0.004	Alfa 3.70	-1.058	0.311	0.004
Alfa 1.71	0.149	0.307	0.012	Alfa 2.71	1.164	0.207	0.002	Alfa 3.71	-1.009	0.191	0.000
Alfa 1.72	-0.950	0.341	0.012	Alfa 2.72	1.089	0.403	0.002	Alfa 3.72	-0.712	0.573	0.000
Alfa 1.73	0.787	0.254	0.013	Alfa 2.73	1.023	0.781	0.002	Alfa 3.73	-1.301	0.344	0.000
Alfa 1.74	0.056	0.085	0.015	Alfa 2.74	2.133	0.122	0.002	Alfa 3.74	-0.914	0.061	0.000
Alfa 1.75	1.272	0.323	0.012	Alfa 2.75	2.535	0.516	0.003	Alfa 3.75	0.760	0.413	0.000
Alfa 1.76	0.464	0.241	0.012	Alfa 2.76	1.395	0.213	0.002	Alfa 3.76	-0.919	0.140	0.002
Alfa 1.77	1.136	0.436	0.012	Alfa 2.77	0.961	0.311	0.002	Alfa 3.77	-0.766	0.067	0.000
Alfa 1.78	-0.247	0.682	0.012	Alfa 2.78	1.186	0.161	0.002	Alfa 3.78	-1.436	0.085	0.001
Alfa 1.79	-1.070	0.368	0.016	Alfa 2.79	0.688	0.071	0.002	Alfa 3.79	-0.983	0.075	0.000

Alfa 1.80	1.174	0.120	0.015	Alfa 2.80	1.508	0.171	0.003	Alfa 3.80	1.079	0.071	0.005
Alfa 1.81	0.991	0.510	0.012	Alfa 2.81	0.921	0.504	0.002	Alfa 3.81	-1.748	0.091	0.002
Alfa 1.82	0.584	0.603	0.015	Alfa 2.82	1.186	0.055	0.003	Alfa 3.82	-1.254	0.056	0.000
Alfa 1.83	0.940	0.278	0.019	Alfa 2.83	0.853	0.067	0.004	Alfa 3.83	-1.149	0.086	0.004
Alfa 1.84	0.766	0.278	0.019	Alfa 2.84	1.126	0.280	0.004	Alfa 3.84	-1.128	0.378	0.000
Alfa 1.85	0.520	0.206	0.012	Alfa 2.85	1.627	0.121	0.003	Alfa 3.85	-1.122	0.410	0.009
Alfa 1.86	0.800	0.237	0.014	Alfa 2.86	0.816	0.267	0.003	Alfa 3.86	-1.062	0.621	0.000
Alfa 1.87	1.562	0.211	0.014	Alfa 2.87	1.085	0.062	0.003	Alfa 3.87	-2.266	0.053	0.000
Alfa 1.88	0.910	0.135	0.015	Alfa 2.88	1.068	0.671	0.003	Alfa 3.88	-1.163	0.164	0.001
Alfa 1.89	-0.912	0.913	0.012	Alfa 2.89	0.948	0.189	0.002	Alfa 3.89	-1.168	0.813	0.000
Alfa 1.90	-0.977	0.341	0.012	Alfa 2.90	1.196	0.283	0.003	Alfa 3.90	-1.146	0.176	0.003
Alfa 1.91	-0.851	0.507	0.012	Alfa 2.91	1.372	0.644	0.002	Alfa 3.91	1.724	0.318	0.006
Alfa 1.92	0.491	0.106	0.016	Alfa 2.92	0.759	0.050	0.003	Alfa 3.92	-1.163	0.094	0.003
Alfa 1.93	-0.940	0.409	0.013	Alfa 2.93	0.544	0.205	0.003	Alfa 3.93	-1.352	0.071	0.002
Alfa 1.94	-0.048	0.403	0.012	Alfa 2.94	1.191	0.233	0.003	Alfa 3.94	-1.947	0.296	0.000
Alfa 1.95	0.595	0.167	0.012	Alfa 2.95	2.480	0.143	0.002	Alfa 3.95	-1.269	0.083	0.001
Alfa 1.96	0.482	0.408	0.013	Alfa 2.96	1.149	0.425	0.003	Alfa 3.96	-1.137	0.398	0.006
Alfa 1.97	0.248	0.171	0.012	Alfa 2.97	1.212	0.101	0.002	Alfa 3.97	-1.097	0.114	0.000
Alfa 1.98	0.488	0.058	0.013	Alfa 2.98	1.601	0.259	0.003	Alfa 3.98	-1.243	0.154	0.000
Alfa 1.99	0.134	0.839	0.015	Alfa 2.99	0.628	0.071	0.004	Alfa 3.99	-1.249	0.505	0.000
Alfa 1.100	-1.843	0.554	0.011	Alfa 2.100	0.689	0.846	0.002	Alfa 3.100	-0.224	0.208	0.000

ANEXO C - CLV ESTIMADO PARA CADA CLIENTE

Modelo aplicado			Modelo Ingênuo		
Ordem	Cod. Cliente	Valor do cliente	Ordem	Cod. Cliente	Valor do cliente
1	27	47.831,08	1	27	39.072,81
2	47	30.680,82	2	40	26.272,65
3	4	26.713,68	3	93	19.798,00
4	49	16.843,37	4	28	18.879,49
5	30	16.683,78	5	73	14.182,91
6	36	10.819,99	6	88	14.138,09
7	88	7.480,28	7	86	10.447,34
8	91	7.312,45	8	97	9.034,99
9	73	7.282,59	9	63	6.172,21
10	25	6.048,97	10	4	4.270,05
11	90	5.915,76	11	47	4.247,82
12	39	5.797,87	12	17	3.733,66
13	59	5.152,78	13	91	3.704,74
14	45	4.626,90	14	80	3.701,50
15	5	2.869,30	15	65	3.657,87
16	10	2.858,60	16	59	3.548,83
17	62	2.186,15	17	20	3.447,59
18	100	1.947,00	18	81	3.435,88
19	40	1.440,38	19	87	3.382,70
20	70	1.188,11	20	78	3.213,54
21	12	958,14	21	34	3.154,46
22	22	707,79	22	49	3.143,66
23	20	675,67	23	100	1.965,26
24	44	441,36	24	82	1.963,36
25	97	244,61	25	25	1.602,07
26	96	1,45	26	43	1.584,52
27	7	-22,78	27	76	943,83
28	85	-30,26	28	66	768,57
29	34	-52,74	29	15	745,40
30	35	-55,41	30	5	721,54
31	82	-64,74	31	48	711,57
32	77	-70,40	32	90	689,63
33	13	-71,80	33	53	622,80
34	50	-73,14	34	52	553,38
35	93	-76,41	35	69	529,49
36	19	-78,96	36	98	502,40
37	9	-80,57	37	31	465,96
38	26	-81,27	38	13	464,69
39	54	-82,54	39	67	445,94
40	98	-83,69	40	23	441,84

41	37	-84,77	41	79	393,88
42	78	-85,06	42	94	388,16
43	14	-86,00	43	42	367,41
44	24	-86,25	44	32	352,36
45	1	-89,48	45	61	338,95
46	68	-91,25	46	71	328,52
47	11	-91,37	47	75	321,32
48	29	-91,69	48	30	216,53
49	69	-92,52	49	33	172,08
50	17	-92,76	50	36	163,40
51	43	-94,82	51	37	143,03
52	46	-96,27	52	12	137,91
53	33	-99,74	53	24	119,99
54	21	-101,49	54	85	89,33
55	75	-101,86	55	2	88,30
56	89	-103,64	56	50	86,42
57	74	-104,87	57	45	77,84
58	52	-105,50	58	96	50,58
59	95	-105,68	59	29	42,68
60	71	-107,29	60	9	38,84
61	79	-107,51	61	54	36,12
62	18	-108,20	62	18	35,78
63	58	-108,81	63	39	22,91
64	23	-111,78	64	57	18,41
65	60	-116,88	65	26	17,82
66	63	-121,37	66	74	16,60
67	94	-123,77	67	35	11,17
68	31	-124,42	68	64	10,68
69	57	-125,32	69	11	6,50
70	72	-126,05	70	68	5,71
71	2	-126,16	71	19	2,18
72	8	-127,49	72	21	-1,81
73	66	-129,25	73	72	-2,29
74	3	-130,85	74	60	-3,74
75	38	-131,40	75	46	-4,62
76	64	-132,20	76	8	-4,77
77	42	-142,50	77	1	-8,12
78	28	-142,90	78	3	-12,08
79	6	-145,86	79	83	-13,53
80	99	-146,76	80	89	-20,75
81	76	-150,58	81	51	-23,51
82	48	-151,53	82	58	-23,79
83	65	-151,72	83	95	-26,22
84	16	-152,99	84	7	-29,60

85	67	-155,82	85	14	-34,41
86	51	-157,50	86	77	-35,35
87	15	-162,90	87	62	-37,78
88	92	-163,26	88	10	-37,89
89	84	-176,15	89	44	-43,21
90	83	-181,55	90	56	-59,35
91	81	-190,63	91	22	-73,34
92	55	-195,93	92	6	-74,22
93	41	-201,20	93	38	-74,67
94	32	-208,20	94	70	-106,40
95	53	-213,18	95	99	-143,99
96	87	-221,88	96	16	-145,54
97	86	-241,61	97	41	-173,51
98	61	-254,21	98	92	-196,96
99	80	-274,08	99	55	-247,11
100	56	-274,28	100	84	-303,60

ANEXO D - ANÁLISE CLIENTE A CLIENTE PARA VERIFICAÇÃO DA CAPACIDADE DE PREVISÃO

Cód. Cliente	Análise	Status da previsão
1	Realizou compra apenas em 03/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK
2	Realizou compras apenas em 03 e 06/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK
3	Realizou compras apenas em 07 e 08/2012 e não comprou mais. Possui <i>CLV</i> negativo.	OK
4	Este é um ótimo cliente, pois comprou todos os meses desde 03/2011 gerando margens de contribuição média de R\$ 150,00, com <i>net</i> crescente atingindo os R\$ 600.000,00 e com diversos dias de cada mês com transação. <i>CLV</i> de R\$ 26.713,68.	OK
5	Realizou compras em 8 meses desde 06/2012 tendo valor crescente da margem de contribuição e do <i>net</i> . <i>CLV</i> de R\$ 2.869,30	OK
6	Realizou compras em apenas seis meses dos 29 analisados, sendo a última em 01/2013. Possui margem de contribuição baixa e sem tendências de alta. Além disso, zerou o seu <i>net</i> desde a última transação realizada. Possui <i>CLV</i> negativo	OK
7	Cliente realizou transações nos últimos períodos, no entanto, mesmo nos períodos observados, já havia gerado mais custos de <i>marketing</i> do que margem de contribuição. Possui <i>CLV</i> negativo.	OK
8	Realizou apenas uma compra em 03/2013 com margem de contribuição baixa e após zerou o seu <i>net</i> . Possui <i>CLV</i> negativo	OK
9	Realizou compras nos últimos dois períodos, sendo que em um deles teve de R\$ 60,00 de margem de contribuição. É um cliente novo, com frequência de compra baixa e aumentou seu <i>net</i> de R\$ 0,00 para R\$ 19.000,00 nos últimos quatro meses. No entanto as previsões apontam para não compra nos períodos futuros. Parece mais coerente existirem compras no futuro. Possui <i>CLV</i> negativo.	ERRO
10	Realizou compras em 4 dos últimos 7 meses. Possui margem baixa, porém número de dias com transação elevou-se. A frequência de compra é baixa, pois é um cliente novo. A margem de contribuição prevista está um pouco inflada, porém a previsão é aceitável, pois está prevista a realização de compras nos períodos futuros. <i>CLV</i> é R\$ 2.858,60.	OK
11	Realizou compra apenas em 07/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK
12	Realizou compras em 11 dos últimos 14 meses. Possui valor médio de margem em torno de R\$ 15,00. Possui <i>net</i> estabilizado em aproximadamente R\$ 50.000,00. Cliente novo, possui frequência de compra mínima. <i>CLV</i> é R\$ 958,14	OK
13	Realizou compras apenas em 03 e 05/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK
14	Realizou compra apenas em 03 e 04/2011 e em 01/2013. Sendo que nesse último mês a margem de contribuição foi abaixo de R\$ 1,00. A frequência de compra é alta e o <i>net</i> é decrescente, sendo que após a última compra ficou R\$ 15,00. Possui <i>CLV</i> negativo	OK

15	Realizou compras em 14 dos 29 meses, gerando margens de contribuição de até R\$ 326 reais. Chegou a realizar compras em 19 dias em um único mês. O <i>net</i> em 2011 era de cerca de R\$ 13.000,00. Certamente tratava-se de um bom cliente. No entanto foi gradativamente reduzindo seu <i>net</i> , chegando a R\$ 320,00, parou de realizar compras nos últimos períodos e as margens de contribuição caíram. Possui <i>CLV</i> negativo.	OK
16	Trata-se de um caso especial, uma vez que o cliente possuía <i>net</i> de apenas R\$ 100,00 e investia em fundos trazendo margens de contribuição muito baixas até 07/2012, quando aumentou seu <i>net</i> para até R\$ 20.000,00 em 09/2012, investiu dois meses em Bovespa e, nos meses seguintes, reduziu seu <i>net</i> novamente para R\$ 100,00. Diante dessa tendência final de deserção, seu <i>CLV</i> estimado é negativo.	OK
17	Realizou compras apenas nos primeiros meses até 07/2011. Após, praticamente zerou seu <i>net</i> . Possui <i>CLV</i> negativo.	OK
18	De forma semelhante ao cliente 16, houve compras somente de 01 a 03/2013 e o <i>net</i> aumentou de R\$ 0,00 para R\$ 2.500,00. Houve margens de até R\$ 30,00. No entanto, nos dois meses seguintes, o <i>net</i> foi praticamente zerado, sendo que, no último período, não houve compra. Diante dessa tendência final de deserção, seu <i>CLV</i> estimado é negativo.	OK
19	Realizou compras em 02 e 03/2011 e apenas uma compra 02/2013 com uma margem de contribuição de R\$ 4,00. Apesar de haver uma elevação do <i>net</i> para de R\$ 1.000,00 para R\$ 9.000,00, cliente comprou apenas em 3 meses durante os 29 meses, adquiriu apenas uma categoria de produtos no único mês com compra e como se trata de um cliente antigo, sua frequência de compra é de uma compra a cada 10 meses. Diante disso, entende-se haver consistência na previsão de <i>CLV</i> negativo.	OK
20	Realizou compras em 25 dos 29 períodos. Não comprou nos últimos dois meses, no entanto seu <i>net</i> manteve-se constante em cerca de R\$ 50.000,00. Possui margem de contribuição média de R\$ 130,00. Diante dessas não compras nos últimos períodos, a margem prevista acabou penalizada, o que é o único problema que pode ser melhorado na previsão. <i>CLV</i> é R\$ 675,67.	OK
21	Realizou compras apenas nos primeiros meses até 05/2011. Após praticamente zerou seu <i>net</i> . Possui <i>CLV</i> negativo.	OK
22	Realizou compras em 8 dos 29 períodos, sem que nos últimos meses compras fossem realizadas. No entanto esse cliente possui <i>net</i> de R\$ 250,00 e margem de contribuição muito baixas, ficando próximo de R\$ 1,00 em média. Diante disso, entende-se que embora as previsões para os custos de <i>marketing</i> e para a incidência de compra estejam consistentes, o valor das margens de contribuição está muito elevado para o potencial desse cliente. <i>CLV</i> é R\$ 707,79	ERRO
23	Realizou compras em 7 dos 29 períodos. Sua última compra foi em 07/2012 e o <i>net</i> foi gradativamente reduzido até chegar a zero. Possui <i>CLV</i> negativo.	OK
24	Realizou 18 compras nos 29 períodos e seu <i>net</i> está estabilizado em torno de R\$ 20.000. Cliente também comprou no último período. Como esse cliente investe apenas em fundos, possui valor zerado para o total de dias com transação devido à falta dessa informação na base de dados. Diante disso, a previsão para esse cliente está equivocada. Possui <i>CLV</i> negativo.	ERRO
25	Realizou compras em todos os últimos 20 meses com margens de contribuição de até R\$ 415,00. Apesar da redução de seu <i>net</i> nos últimos meses, manteve-se negociando em torno de 4 dias por mês. <i>CLV</i> é R\$ 6048,97.	OK
26	Realizou compra apenas em 03/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK

27	Realizou compras em 25 dos 29 períodos. Possui <i>net</i> em torno de R\$ 1.900.000,00. Possui margem de contribuição média em torno de R\$ 1.400,00. Nos últimos 4 meses, o total de dias com transação subiu chegando a 23 dias com transação em um único mês. <i>CLV</i> é R\$ 47.831,08.	OK
28	Realizou compras em 17 dos 29 períodos. O cliente possuía margens altas em torno de R\$ 1.000,00. No entanto sua última compra foi em 10/2012. Após, não comprou mais e seu <i>net</i> caiu de R\$ 282.000,00 no início de 2011 para R\$ 389,00 em 05/2013. Claramente, é um cliente que desertou. Possui <i>CLV</i> negativo.	OK
29	Realizou compra apenas em 10/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK
30	Realizou compras apenas nos últimos 4 meses. No entanto as margens de contribuição chegaram a R\$ 140,00. Seu <i>net</i> aumentou de R\$ 0,00 em 01/2013 para R\$ 20.000,00 em 05/2013. Como é um cliente novo, sua frequência de compra é a mínima possível. Além disso, chegou a comprar em 14 dias em um único mês. <i>CLV</i> é R\$ 16.683,78.	OK
31	Realizou compras em apenas 3 dos 29 períodos, sendo que sua compra foi em 09/2012. Possui <i>CLV</i> negativo.	OK
32	Realizou compras em apenas 12 dos 29 períodos. O cliente comprava frequentemente, apesar de não gerar margens de contribuição altas. No entanto sua última compra foi em 07/2012. Após, não comprou mais e seu <i>net</i> caiu de para R\$ 0,00. Claramente, é um cliente que desertou. Possui <i>CLV</i> negativo.	OK
33	Realizou compras em 21 dos 29 períodos. No entanto, desde 02/2012, a margem de contribuição média tem ficado em torno de R\$ 1,20 e seu <i>net</i> reduziu de R\$ 41.000,00 para R\$ 3.900,00. Além disso, como esse cliente investe apenas em fundos, não há informação sobre o total de dias com transação. Diante disso, apesar de o cliente ter comprado nos últimos meses e diversas compras tenham sido previstas, os custos de <i>marketing</i> são superiores. Possui <i>CLV</i> é negativo.	OK
34	Realizou compras em 06 e 11/2012. Sendo que em 06/2012 a margem de contribuição foi de R\$ 3.000,00. Após, além de ter deixado de comprar, seu <i>net</i> foi zerado. É mais um cliente que desertou. Possui <i>CLV</i> negativo.	OK
35	Realizou compra apenas em 05/2011 e não comprou mais. Possui <i>CLV</i> negativo.	OK
36	Realizou apenas 4 compras desde 10/2012. Como é um cliente novo, possui frequência de compra baixa de uma compra a cada 2 meses. Além disso, esse cliente aumentou seu <i>net</i> de R\$ 0,00 para R\$ 1.145.000,00. Diante disso, apesar de investir apenas em fundos e de o total de períodos com transação não estar disponível, é perfeitamente coerente esperar um valor futuro elevado para esse cliente. <i>CLV</i> é R\$ 10.819,99.	OK
37	Realizou compras em apenas 3 dos 29 períodos, sendo que sua compra foi em 07/2011. Possui <i>CLV</i> negativo.	OK
38	Realizou compras em 8 dos 29 períodos. O <i>net</i> aumentou de R\$ 19.000 para R\$ 30.000,00. Apesar desse <i>net</i> relevante, as margens desse cliente são muito baixas, em torno de R\$ 3,00. Mesmo nos períodos observados, os custos de <i>marketing</i> já haviam sido superiores em relação às margens de contribuição geradas. Sendo assim, o <i>CLV</i> negativo estimado é coerente.	OK
39	Realizou compras em 5 dos 29 períodos. Como é um cliente novo, possui frequência de compra baixa de uma compra a cada 2 meses. Além disso, esse cliente aumentou seu <i>net</i> de R\$ 0,00 para quase R\$ 60.000,00. Apesar de o valor futuro estar um pouco inflado, o fato de o potencial de crescimento ter sido captado pelo modelo torna a previsão aceitável. <i>CLV</i> é R\$ 5797,87.	OK
40	Apesar desse cliente ter tido margens de contribuição de até R\$ 8.800,00 até o final de 2012, em 2013, nenhuma compra foi	ERRO

	realizada e o <i>net</i> desse cliente foi zerado nos últimos períodos. Sendo assim, o <i>CLV</i> de R\$ 1440,38 não parece condizer com a provável condição de deserção desse cliente.	
41	Realizou compras em 11 dos 29 períodos. O <i>net</i> diminuiu de R\$ 17.000 para R\$ 6.400,00. Além disso, as margens desse cliente são muito baixas, em torno de R\$ 0,50. Mesmo nos períodos observados, os custos de <i>marketing</i> já haviam sido bastante superiores em relação às margens de contribuição geradas. Sendo assim, apesar de o total de dias com transação ser 0, porque se trata de um cliente de fundos, o <i>CLV</i> negativo estimado é coerente.	OK
42	Realizou compras em 13 dos 29 períodos. No entanto sua última compra foi em 10/2012. Após, além de não comprar, seu <i>net</i> foi reduzido para R\$ 14,00. Trata-se de outro cliente que desertou. Possui <i>CLV</i> negativo.	OK
43	Realizou compras em 6 dos 29 períodos. No entanto sua última compra foi em 11/2011. Após, além de não comprar, seu <i>net</i> foi reduzido para R\$ 0,00. Trata-se um cliente que desertou. Possui <i>CLV</i> negativo.	OK
44	Realizou compras em 3 dos 29 períodos, sendo uma delas no último período observado. É um cliente com <i>net</i> de apenas R\$ 800,00. Além disso, as margens desse cliente são muito baixas, em torno de R\$ 0,40. Mesmo nos períodos observados, os custos de <i>marketing</i> já haviam sido bastante superiores em relação às margens de contribuição geradas. Sendo assim, mesmo com o total de dias com transação ser 0, porque se trata de um cliente de fundos, o <i>CLV</i> de R\$ 441,36 não é consistente.	ERRO
45	Realizou compras em 13 dos 29 períodos durante todo o período. Sendo assim, sua frequência de compra é baixa. Suas margens de contribuição possuem valor médio em torno de até R\$ 50,00. O <i>net</i> manteve-se em aproximadamente R\$ 20.000 mesmo no último período. Sendo assim, o <i>CLV</i> de R\$ 4.626,90.	OK
46	Realizou compras em apenas 2 dos 29 períodos, sendo que sua compra foi em 09/2012. Possui <i>CLV</i> negativo.	OK
47	Realizou apenas 2 compras nos últimos 4 períodos. No entanto, como trata-se de um cliente novo, sua frequência de compra é de uma compra a cada 2 meses. Além disso, no penúltimo mês, sua margem de contribuição foi de R\$ 4042,70 e o seu <i>net</i> subiu de 0 para R\$ 239.360,38. Sendo assim o alto potencial de rentabilidade desse cliente é traduzido no valor de <i>CLV</i> de R\$ 30.680,82.	OK
48	Realizou compras em apenas 3 dos 29 períodos, sendo que sua última compra foi em 05/2011. Possui <i>CLV</i> negativo.	OK
49	Realizou compras em 9 dos 29 períodos, sendo que seis deles nos últimos períodos, quando a margem de contribuição de um desses meses chegou a R\$ 2.547,68 em um dos meses. O <i>net</i> também elevou-se de cerca de \$ 20.000,00 para pouco mais de R\$ 260.000,00. <i>CLV</i> é R\$ 16.843,37.	OK
50	Esta é uma previsão difícil de ser realizada. Trata-se de um cliente novo que realizou 3 compras desde 10/2012. Suas margens de contribuição foram no máximo de R\$ 50,00. No entanto seu <i>net</i> elevou para R\$ 305.000,00. Como se trata de um cliente que comprou fundos, não há informação sobre o total de dias com transação. Além disso, não houve compras nos últimos dois períodos. Diante disso, seu <i>CLV</i> é negativo. Frente a essa situação, entende-se que o potencial do cliente traduzido pelo aumento do seu <i>net</i> não foi capturado.	ERRO
51	Realizou compras em apenas 3 dos 29 períodos, sendo que sua última compra foi em 06/2012. Possui <i>CLV</i> negativo.	OK
52	Este é um bom cliente, com compras em apenas 18 dos 29 períodos, com margem média de R\$ 30,00 e <i>net</i> crescente. No entanto o modelo não foi capaz de captar esse panorama. Possui <i>CLV</i> negativo.	ERRO

53	Realizou compras em apenas 10 dos 29 períodos, sendo que o último período com compra foi 01/2011. Seu <i>net</i> manteve-se constante em cerca de R\$ 13.000,00. Apesar da dificuldade de prever os fluxos de caixa futuros desse cliente, entende-se que um valor negativo de -241,18 para o seu <i>CLV</i> não condiz com o seu potencial.	ERRO
54	Realizou compras em 2 dos 29 períodos. Apesar de ser um cliente novo, sua última compra com margem de contribuição alta de R\$ 185,00 foi em 10/2012. Trata-se de um cliente que mantém seu <i>net</i> de cerca de R\$ 40.000,00, porém não transaciona. Sendo assim entende-se que seu <i>CLV</i> negativo de R\$ 82,54 é coerente.	OK
55	Realizou compras em 20 dos 29 períodos. No entanto suas margens de contribuição não passam de R\$ 2,50 e sua última compra foi em 11/2012. Após, seu <i>net</i> foi derrubado para zero. Possui <i>CLV</i> negativo.	OK
56	Realizou compras em 16 dos 29 períodos. No entanto sua última compra foi em 12/2012. Após, seu <i>net</i> foi derrubado para zero. Possui <i>CLV</i> negativo.	OK
57	Realizou compras apenas em 05 e 06/2011. Possui <i>CLV</i> negativo.	OK
58	Realizou compras apenas em 12/2011 e 07/2012. Possui <i>CLV</i> negativo.	OK
59	Realizou compras em 9 dos 29 períodos. No entanto trata-se de um cliente novo, portanto sua de compra é baixa. Além disso, seu <i>net</i> aumentou de R\$ 0,00 para quase R\$ 155.000,00. Além disso, obteve margem de até R\$ 3.300,00 em um dos períodos. <i>CLV</i> é R\$ 5152,78.	OK
60	Realizou compras apenas nos primeiros três meses de 2011. Possui <i>CLV</i> negativo.	OK
61	Realizou compras em 24 dos 29 períodos. Apesar do valor baixo de seu <i>net</i> de cerca de R\$ 7.000,00 durante todo o período observado, esse cliente possui uma margem de contribuição de cerca de R\$ 20,00 por mês e ao menos incidência de compra deveriam ter sido previstas, o que não ocorreu. Possui <i>CLV</i> negativo.	ERRO
62	Realizou compras em 19 dos 29 períodos, inclusive nos últimos períodos. Seu <i>net</i> aumentou de R\$ 70.000,00 para cerca de R\$ 108.000,00. <i>CLV</i> é R\$ 2.186,15.	OK
63	Realizou compra apenas de 02 a 04/2013. Possui <i>CLV</i> negativo.	OK
64	Realizou compras apenas em 02 e 03/2013, porém com margens abaixo de R\$ 9,00. Apesar de seu <i>net</i> crescente, não há como ter maiores expectativas em relação a esse cliente. Possui <i>CLV</i> negativo.	OK
65	Realizou compras em 6 dos 29 períodos, sendo que, após 07/2012, quando gerou uma alta margem e contribuição de cerca de R\$ 2600,00, seu <i>net</i> que era de aproximadamente R\$ 120.000,00 foi reduzido para quase 400,00. Possui <i>CLV</i> negativo.	OK
66	Realizou compras em 10 dos 29 períodos, sendo 8 deles até 12/2011. Após, as 2 compras em 07/2012 e 04/2013 foram realizadas para reduzir seu <i>net</i> de cerca de R\$ 19.000, 00 para quase R\$ 150,00. Cliente provavelmente desertou. Possui <i>CLV</i> negativo.	OK
67	Realizou compras em 14 dos 29 períodos, sendo 10 deles até 12/2011. Após, nos outros 4 períodos, foram realizadas transações para reduzir seu <i>net</i> de cerca de R\$ 70.000, 00 para quase R\$ 100,00. Cliente provavelmente desertou. Possui <i>CLV</i> negativo.	OK
68	Realizou apenas uma transação em 03/2013 com margem de R\$ 14,00. Sendo assim, sua previsão não apresenta nenhuma incidência de compra. Possui <i>CLV</i> negativo.	OK
69	Cliente novo e apresentou compras nos três últimos períodos, com margens de até R\$ 424,62. No entanto a sua margem do último	ERRO

	período foi de R\$ 10,00 e o seu <i>net</i> é de apenas cerca de R\$ 1.200,00. Apesar da previsão de incidência de compra para todos os períodos futuros, a previsão da margem de contribuição é zero, o que é inconsistente em relação ao histórico desse cliente. Possui <i>CLV</i> negativo.	
70	Realizou compras em 9 dos 29 períodos. Trata-se de um cliente novo e possui transações nos últimos períodos. Possui também o <i>net</i> com tendência de crescimento. Apesar de ser um cliente de fundos e não ter disponível o valor do total de dias com transação, o <i>CLV</i> é de R\$ 1188,11.	OK
71	Realizou compras em 8 dos 29 períodos. No entanto sua última compra ocorreu em 10/2011. Possui <i>CLV</i> negativo.	OK
72	Realizou compras apenas em 09/2012. Possui <i>CLV</i> negativo.	OK
73	Realizou compras em todos os períodos desde 01/2012. Possui margem de contribuição média de R\$ 811,51. Seu <i>net</i> se manteve estável em torno de R\$ 20.000,00. Além disso, o cliente transaciona mensalmente, tanto na categoria de produtos BMF quanto na categoria de produtos Bovespa. <i>CLV</i> é R\$ 7.282,59.	OK
74	Realizou compras apenas em 04/2013 com margem de contribuição de R\$ 20,00. Possui <i>CLV</i> negativo.	OK
75	Realizou compras apenas em 06/2012. Possui <i>CLV</i> negativo.	OK
76	Realizou compras em 12 dos 29 períodos. No entanto sua última compra ocorreu em 07/2012. Possui <i>CLV</i> negativo.	OK
77	Realizou compras apenas em 10/2012 e 01/2013 com margem de contribuição de, no máximo, R\$ 11,46. Possui <i>CLV</i> negativo.	OK
78	Realizou compras em 6 dos 29 períodos. No entanto sua última compra ocorreu em 09/2011. Possui <i>CLV</i> negativo.	OK
79	Realizou compras apenas em 01 e 02/2011. Possui <i>CLV</i> negativo.	OK
80	Realizou compras em todos os períodos desde 09/2011. Possui margem de contribuição média de R\$ 140,89. Seu <i>net</i> se manteve acima de R\$ 300.000,00 Os únicos pontos negativos são o total de dias com transação por mês que reduziu sucessivamente e o número de categorias de produtos transacionadas, que caiu de 2 para 1. No entanto, mesmo com esses problemas, o <i>CLV</i> negativo de R\$ 274,08 não condiz com o histórico desse cliente.	ERRO
81	Esse cliente investia em BMF e em Bovespa até 03/2012. Até esse período, grandes margens de até R\$ 690,79 foram registradas. No entanto, desse mês em diante, o cliente passou a investir apenas em fundos gerando margens muito baixas, que não superam R\$ 0,20. Além disso, o seu <i>net</i> também decresceu. Diante disso, o <i>CLV</i> negativo estimado é consistente, pois as incidências de compra foram previstas, porém as margens previstas foram zero.	OK
82	Cliente possui 10 períodos com compras nos últimos 13 períodos. Seu <i>net</i> foi mantido em cerca de R\$ 30.000,00 e sua margem de contribuição média foi de R\$ 74,54 havendo uma considerável queda das margens nos últimos períodos com transação. O total de dias com transação também decresceu muito de 15 para até 3 dias. Apesar de pontos negativos, incidências de compra futuras e margens de contribuição acima de zero deveriam ser observadas. Possui <i>CLV</i> negativo.	ERRO
83	Realizou compras em 20 dos 29 períodos. No entanto sua margem de contribuição média é apenas R\$ 8,32. Como o cliente deixou de investir em Bovespa e passou a investir em fundos, o total de dias com transação decresceu para zero. O <i>net</i> manteve-se em aproximadamente R\$ 20.000,00. Sendo assim, embora o cliente tenha transacionado menos nos últimos períodos, seria plausível ter	ERRO

	incidências de compras e margens de contribuições previstas. Possui <i>CLV</i> negativo.	
84	Realizou compras em 27 dos 29 períodos. O <i>net</i> manteve-se em apenas cerca de R\$ 800,00 durante todo o período observado. Além disso, as margens desse cliente são muito baixas em torno de R\$ 0,50. Mesmo nos períodos observados os custos de <i>marketing</i> já haviam sido bastante superiores em relação às margens de contribuição geradas. Sendo assim, apesar de o total de dias com transação ser 0, porque se trata de um cliente de fundos, o <i>CLV</i> negativo estimado é coerente.	OK
85	Realizou apenas 4 compras nos últimos períodos. No entanto trata-se de um cliente novo e seu <i>net</i> aumentou para cerca de R\$ 300.000,00. Sendo assim, apesar de as margens ainda não serem elevadas, o cliente apresenta grande potencial que não foi captado pelo modelo. Possui <i>CLV</i> negativo.	ERRO
86	Realizou compras em 18 dos 29 períodos. No entanto sua última compra ocorreu em 10/2012. Após, seu <i>net</i> foi praticamente zerado. Possui <i>CLV</i> negativo.	OK
87	Realizou compras em 19 dos 29 períodos. No entanto sua última compra ocorreu em 7/2012. Após, seu <i>net</i> foi praticamente zerado. Possui <i>CLV</i> negativo.	OK
88	Realizou compras em 18 dos 29 períodos, todas nos últimos períodos. Sua margem de contribuição média é R\$ 765,37. <i>CLV</i> é 7.480,28.	OK
89	Realizou compras apenas em 01/2013 com margem de contribuição de R\$ 3,06. Possui <i>CLV</i> negativo.	OK
90	Realizou 3 compras nos últimos 5 períodos. Trata-se de um cliente novo, portanto a frequência de compra é baixa. Seu <i>net</i> é de cerca de R\$ 90.000,00 e teve margem de contribuição de até R\$ 642,64. <i>CLV</i> é 5.915,76.	OK
91	Realizou compras em 17 dos 29 períodos, sendo que 7 deles são os últimos períodos. Sua margem de contribuição média chegou a R\$ 876,78. <i>CLV</i> é 7.312,45.	OK
92	Realizou compras em 17 dos 29 períodos. No entanto sua última compra ocorreu em 10/2012. Após, seu <i>net</i> foi zerado. Possui <i>CLV</i> negativo.	OK
93	Realizou compras em 9 dos 29 períodos. No entanto sua última compra ocorreu em 03/2012. Após, seu <i>net</i> foi zerado. Possui <i>CLV</i> negativo.	OK
94	Realizou compras em 7 dos 29 períodos. No entanto sua última compra ocorreu em 05/2012. Após, seu <i>net</i> foi zerado. Possui <i>CLV</i> negativo.	OK
95	Realizou compras apenas em 01 e 02/2013 com margens de contribuição menores do que R\$ 0,03. Possui <i>CLV</i> negativo.	OK
96	Realizou compras em 9 dos 29 períodos, todas nos últimos períodos. Seu <i>net</i> manteve-se em aproximadamente R\$ 50.000,00. Sua margem de contribuição média é de R\$ 19,57. <i>CLV</i> é apenas R\$ 1,45, porém, poderia ser mais elevado.	ERRO
97	Realizou compras em 9 dos 29 períodos. No entanto deixou de comprar nos últimos três períodos, e seu <i>net</i> diminuiu bastante. Sua margem de contribuição média é de R\$ 975,78. <i>CLV</i> é R\$ 244,61. Deveria ser negativo, pois o cliente provavelmente se tornará inativo.	ERRO
98	Realizou compras apenas de 02 a 06/2011. Possui <i>CLV</i> negativo.	OK
99	Realizou compras em 17 dos 29 períodos. No entanto sua última compra ocorreu em 09/2012. Após, seu <i>net</i> foi zerado. Possui <i>CLV</i>	OK

	negativo.	
100	Realizou compras em 27 dos 29 períodos. Sua margem de contribuição média é de R\$ 80,60. Seu <i>net</i> é de cerca de R\$ 190.000,00. <i>CLV</i> é R\$ 1947,00	OK