UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

ESTUDOS EMPÍRICOS SOBRE MICROECONOMIA BANCÁRIA NO BRASIL

Celina Yumiko Ozawa

Orientador: Prof. Dr. Márcio Issao Nakane

SÃO PAULO

2008

Profa. Dra. Suely Vilela
Reitora da Universidade de São Paulo

Prof. Dr. Carlos Roberto Azzoni
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Joaquim Martins Guilhoto
Chefe do Departamento de Economia

Prof. Dr. Dante Mendes Aldrighi
Coordenador do Programa de Pós-Graduação em Economia

CELINA YUMIKO OZAWA

ESTUDOS EMPÍRICOS SOBRE MICROECONOMIA BANCÁRIA NO BRASIL

Tese apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito para a obtenção do título de Doutora em Economia.

Orientador: Prof. Dr. Márcio Issao Nakane

SÃO PAULO

2008

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Ozawa, Celina Yumiko

Estudos empíricos sobre microeconomia bancária no
Brasil / Celina Yumiko Ozawa. – São Paulo, 2008.
165 p.

Tese (Doutorado) – Universidade de São Paulo, 2008
Bibliografia

1. Bancos (Economia) 2. Crédito 3. Econometria I.
I. Universidade de São Paulo. Faculdade de Economia,
Administração e Contabilidade. II. Título.

CDD – 332.1

AGRADECIMENTOS

Em primeiro lugar ao professor e orientador Márcio Issao Nakane por todo apoio, paciência e dedicação recebidos continuamente durante todo o período da pesquisa. Não poderia ter melhor orientação que a recebida. Muito obrigada por compartilhar da sua sabedoria e da sua experiência acadêmica.

Ao professor Naércio Aquino Menezes Filho, membro da banca de qualificação, agradeço as críticas e os comentários que muito contribuíram para o desenvolvimento deste trabalho.

Ao professor Gabriel de Abreu Madeira pelas valiosas críticas e sugestões recebidas no exame de qualificação e pela atenção e presteza em fornecer sempre um excelente atendimento.

Ao Sérgio Mikio Koyama por disponibilizar os dados, essenciais para a concretização desta tese. Muito obrigada.

Ao professor Fernando Antonio Slaibe Postali pela amizade e incentivo. O mundo precisa de mais professores como você.

Aos funcionários da biblioteca pelo ótimo atendimento. À Márcia, Camila e à Coordenação de Pós-Graduação pela atenção e apoio recebidos. À Valéria e a todos da Secretaria de Pós-Graduação muito obrigada pela atenção e gentileza com que tratam a todos. Todos vocês são o meio que permitem que o nosso trabalho seja concluído.

Ao CNPQ pelo apoio financeiro.

À Fernanda, Henrique e Alexandre pela amizade, inúmeras conversas e dúvidas.

À Aki e Ernesto pelo companheirismo, compreensão, inestimável amizade e por tudo.

A todos que direta ou indiretamente contribuíram para a conclusão deste trabalho e que injustamente não foram citados.

RESUMO

Este estudo está organizado em três capítulos, independentes entre si e que possuem em comum o fato de serem trabalhos aplicados de microeconometria bancária e de utilizarem dados desagregados do Banco Central do Brasil. Os dois primeiros estão relacionados com crédito e o terceiro com os submercados de agências bancárias. Os determinantes da demanda por financiamento de automóveis é o tema central do primeiro capítulo. Usando dados da Central de Risco de Crédito do Banco Central, estima-se as elasticidades da demanda por empréstimo com respeito à taxa de juros e à duração. Os tomadores respondem tanto a uma variável quanto a outra, indicando que não somente a duração, mas a taxa de juros também é importante na decisão de financiamento de automóveis. O segundo capítulo também usa os dados da Central de Risco (na modalidade capital de giro) e trata da relação entre a taxa de juros e a garantia, e da importância do tempo de relacionamento nos contratos de crédito. As estimativas fornecem evidências de que a taxa de juros e a garantia são negativamente relacionadas, e este resultado é consistente com a literatura que dá suporte ao *sorting-by-private information paradigm*. Adicionalmente, um tempo de relacionamento mais longo contribui para mitigar os efeitos da assimetria informacional. No terceiro capítulo, testa-se as predições do modelo de submercados independentes (Sutton, 1998) usando dados de agências bancárias. De acordo com este modelo, o equilíbrio para o agregado de submercados deve ocorrer com um grau mínimo de desigualdade no tamanho das firmas. Quando se emprega os dados do mercado bancário nacional verifica-se que as firmas apresentam alto grau de desigualdade, mas quando se utiliza dados de mercados territorialmente desagregados (microrregiões e municípios), o grau de desigualdade torna-se baixo. Esses resultados estão em consonância com o modelo teórico.

ABSTRACT

This study is organized in the format of three independent papers that have in common: a) the fact that they are all studies about microeconomics of banking applied to Brazil; and b) the use of disaggregated data from Brazilian Central Bank. The first two are related to banking credit and the third one with banking submarkets. The determinants of credit demand on auto loan contracts is the central theme of the first chapter. Using data from the Central Bank's Credit Bureau, we estimated the elasticities of loan demand with respect to interest rate and maturity. Borrowers are responsive to both variables, indicating that loan maturity is not the only important variable in car financing decisions. The second chapter investigates the relationship between interest rates and collateral and the role of relationship lending in the loan contracts, also using data from the Central Bank's Credit Bureau (working capital finance). We find that interest rates and collateral are negatively related, and this result is consistent with the literature that provides support to sorting-by-private information paradigm. Also, a long-term relationship with the lender contributes to mitigate the effects of information asymmetry. In the third chapter, we test the predictions of the independent submarket model (Sutton, 1998) with banking branch level data. According to this model, the equilibrium outcome for aggregates of submarkets must involve some minimal degree of inequality in the size of firms. When using data from the national bank market we find that the degree of inequality of firms is high but when using regionally disaggregated markets (micro regions and municipalities), the degree of inequality becomes low. These findings are in consonant agreement with the theoretical model.

SUMÁRIO

LISTA DE TABELAS.....	4
LISTA DE FIGURAS.....	6
APRESENTAÇÃO	7
1 MERCADO DE CRÉDITO NO BRASIL: UM ESTUDO SOBRE OS IMPACTOS DA TAXA DE JUROS E DO PRAZO DE PAGAMENTO NOS FINANCIAMENTO DE VEÍCULOS PARA PESSOAS FÍSICAS.....	10
1.1 INTRODUÇÃO	10
1.2 BREVE DESCRIÇÃO DO MERCADO DE FINANCIAMENTO DE VEÍCULOS NO BRASIL	14
1.3 ARCABOUÇO TEÓRICO	19
1.3.1 Modelo Simples.....	20
1.4 MODELO EMPÍRICO	25
1.4.1 Seleção da amostra	25
1.4.2 Variáveis endógenas	26
1.4.3 Especificação econométrica	28
1.5 BASE DE DADOS	31
1.5.1 Análise preliminar dos dados	33
1.6 RESULTADOS	37
1.6.1 Amostra completa.....	37
1.6.2 Subdivisão da amostra	42
1.7 CONCLUSÃO.....	46
1.8 REFERÊNCIAS BIBLIOGRÁFICAS	48
APÊNDICE 1.A	50
APÊNDICE 1.B.....	52
APÊNDICE 1.C	55
2 UM ESTUDO SOBRE A RELAÇÃO ENTRE TAXAS DE JUROS, GARANTIAS E TEMPO DE RELACIONAMENTO NOS CONTRATOS DE CRÉDITO	56
2.1 INTRODUÇÃO	56
2.2 HIPÓTESES EMPÍRICAS E REVISÃO DA LITERATURA	59
2.2.1 Relação entre taxa de juros e garantia	59
2.2.2 Tempo de relacionamento entre o tomador e o banco.....	61

2.3	ESTIMAÇÃO	63
2.3.1	Equação da garantia	63
2.3.2	Equação da taxa de juros.....	64
2.3.2.1	<i>Matching Estimator</i>	65
2.3.2.2	<i>Average Treatment Effect</i>	68
2.3.2.3	<i>Switching Regression Model</i>	70
2.3.3	Instrumentos.....	72
2.4	DADOS.....	77
2.4.1	Base de dados.....	78
2.4.2	Análise preliminar dos dados	79
2.5	RESULTADOS EMPÍRICOS	83
2.5.1	Equação da garantia	83
2.5.2	Equação da taxa de juros.....	87
2.5.3	<i>Treatment effects</i>	91
2.5.4	<i>Switching regression model</i>	95
2.5.5	Estimações adicionais	96
2.6	CONCLUSÃO	99
2.7	REFERÊNCIAS BIBLIOGRÁFICAS.....	101
APÊNDICE 2.A		104
APÊNDICE 2.B		105
APÊNDICE 2.C		106
APÊNDICE 2.D		108
3	MODELO DE SUBMERCADOS INDEPENDENTES APLICADO AO SETOR BANCÁRIO BRASILEIRO.....	112
3.1	INTRODUÇÃO	112
3.2	MODELO DE SUBMERCADOS INDEPENDENTES.....	118
3.2.1	Modelo de Submercados Independentes.....	118
3.2.2	Definição do mercado	121
3.2.3	Alguns resultados empíricos	123
3.2.4	Modelo aplicado ao mercado bancário	124
3.3	BASE DE DADOS E RESULTADOS EMPÍRICOS.....	126
3.3.1	Base de dados.....	126
3.3.2	Curvas de <i>Lorenz</i>	127
3.3.3	Identificando o tamanho dos submercados	132
3.3.3.1	Identificando submercados na região CO	139
3.3.4	Identificando submercados independentes em municípios com mais de um submercado	142
3.3.5	Mapas dos submercados de agências bancárias	146
3.3.6	Tamanho das firmas	148
3.4	CONCLUSÃO	151

3.5	REFERÊNCIAS BIBLIOGRÁFICAS	153
APÊNDICE 3.A		155
APÊNDICE 3.B		158
APÊNDICE 3.C		160
APÊNDICE 3.D		161
APÊNDICE 3.E		164

LISTA DE TABELAS

Tabela 1.1 – Emplacamento de automóveis e comerciais leves	14
Tabela 1.2 – Aquisição de veículos	17
Tabela 1.3 – Faixas de atraso – Aquisição de veículos – Pessoa Física – R\$ milhões	18
Tabela 1.4 – Estatística descritiva das variáveis	34
Tabela 1.5 – Estatística descritiva das variáveis por quartis de valores de empréstimo	34
Tabela 1.6 – Descrição das variáveis	35
Tabela 1.7 – Resultados do critério de <i>cross-validation</i>	38
Tabela 1.8 – Resultados da estimativa da Equação 1	39
Tabela 1.9 – Elasticidade da demanda por financiamento de veículos – Modelo 9	41
Tabela 1.10 – Elasticidades para o Modelo 9 para as diversas subamostras	43
Tabela 2.1 – Descrição dos instrumentos.....	75
Tabela 2.2 – Descrição das variáveis	79
Tabela 2.3 – Estatística descritiva das variáveis – Capital de Giro	80
Tabela 2.4 – Estatística descritiva das variáveis por garantia – Capital de Giro	82
Tabela 2.5 – Resultados das estimativas do modelo <i>probit</i>	84
Tabela 2.6 – Efeitos marginais do modelo <i>probit</i>	85
Tabela 2.7 – Resultados das estimativas da equação da taxa de juros – OLS e 2SLS.....	88
Tabela 2.8 – Resultados das estimativas do <i>matching estimator</i>	92
Tabela 2.9 – Resultados das estimativas do modelo de efeitos do tratamento	93
Tabela 2.10 – Resultado do <i>endogenous switching regression model</i>	96
Tabela 3.1 – Número de instituições do segmento bancário segundo porte por ativos totais - 1996-2006	113
Tabela 3.2 – Participação percentual nos ativos totais do segmento bancário - 1996-2006..	113
Tabela 3.3 – Índices de Concentração (C3, C5 e C10) - 1994-2006 – conglomerados bancários	114
Tabela 3.4 – Descrição das variáveis	135
Tabela 3.5 – Número de bancos.....	135
Tabela 3.6 – Estatística descritiva das variáveis	136
Tabela 3.7 – Resultados das estimativas do modelo <i>logit</i> ordenado.....	137
Tabela 3.8 – Variação marginal de <i>nbanks*</i>	138
Tabela 3.9 – Estatística descritiva das variáveis e número de bancos - CO	141
Tabela 3.10 – Resultados das estimativas do modelo <i>logit</i> ordenado para.....	141
Tabela 3.11 – Dados dos municípios selecionados	143
Tabela 3.12 – Índice de Duda e Hart	144

Apêndices:

Tabela A - 1.1 – Parâmetros da distribuição <i>lognormal</i> da Renda	50
Tabela A - 1.2 – Elasticidade da demanda por financiamento de veículos – Modelo 9	55
Tabela A - 1.3 – Elasticidades para o Modelo 9 as diversas subamostras	55
Tabela A - 2.1 – Medidas de eficiência do judiciário por UF - 2005	104
Tabela A - 2.2 – Resultado da equação de seleção do modelo de efeitos do tratamento.....	105
Tabela A - 2.3 – Resultado da equação de seleção do <i>endogenous switching regression model</i>	105
Tabela A - 2.4 – Resultados das estimativas do modelo <i>probit</i> – sem a variável <i>duração</i> ...	106

Tabela A - 2.5 – Resultados das estimativas da equação da taxa de juros – OLS e 2SLS – sem a variável <i>duração</i>	107
Tabela A - 2.6 – Estatística descritiva das variáveis – Capital de Giro – Base de dados reduzida	108
Tabela A - 2.7 – Estatística descritiva das variáveis por garantia – Capital de Giro – Base de dados reduzida	108
Tabela A - 2.8 – Resultados das estimativas do modelo <i>probit</i> – Base de dados reduzida....	109
Tabela A - 2.9 – Efeitos marginais do modelo <i>probit</i> – Base de dados reduzida.....	109
Tabela A - 2.10 – Resultados das estimativas da equação da taxa de juros – OLS e 2SLS – Base de dados reduzida.....	110
Tabela A - 2.11 Resultados das estimativas do <i>matching estimator</i> – Base de dados reduzida.	110
Tabela A - 2.12 – Resultados das estimativas do modelo de efeitos do tratamento – Base de dados reduzida	111
Tabela A - 2.13 – Resultado do <i>endogenous switching regression model</i> – Base de dados reduzida	111
Tabela A - 3.1 – Resultados da estimativa do <i>generalized ordered logit model</i> – municípios (2).....	158
Tabela A - 3.2 – Resultados da estimativa do <i>generalized ordered logit model</i> – municípios (3).....	159

LISTA DE FIGURAS

Figura 1.1– Volume de operações de crédito – Pessoa Física – R\$ milhões.....	15
Figura 1.2 – Taxa de juros pré-fixadas – PF e Selic acumulada no mês (%a.a.).....	16
Figura 1.3 – Forma de pagamento na venda de veículos	17
Figura 1.4: Restrição orçamentária do consumidor que compra e financia.....	23
Figura 2.1 – Volume de operações de crédito (juros pré-fixados) – PJ – R\$ milhões.....	77
Figura 2.2 – Taxa de juros pré-fixadas – PJ e Selic acumulada no mês (% a.a.).....	78
Figura 3.1 – Concentração dos três maiores bancos (C3) – Média de 1995-99	115
Figura 3.2 – Curva de Lorenz do mercado bancário – Brasil	128
Figura 3.3 – Curva de Lorenz para UF - depósito (esquerda superior), crédito (direita superior) e agências (esquerda inferior).....	129
Figura 3.4 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) –Mesorregiões de Pernambuco	130
Figura 3.5 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Municípios com dois ou mais bancos e até 10 mil habitantes	131
Figura 3.6 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Municípios com mais de 500 mil habitantes.....	131
Figura 3.7 – Dendograma para a análise de cluster de Campo Grande – <i>complete linkage</i> (esquerda) e <i>average linkage</i> (direita)	143
Figura 3.8 – Matriz com as observações identificadas pelo número do cluster (<i>average linkage</i>) –.....	144
Figura 3.9 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Campo Grande - MS	145
Figura 3.10 – Mapas de submercados – Brasil (esquerda superior), UFs (direita superior), mesorregiões (esquerda inferior) e microrregiões (direita inferior).....	147
Figura 3.11 – Histograma da participação dos bancos no mercado de Depósito segundo a cobertura geográfica (Nacional ou Local)	150

Apêndices:

Figura A - 3.1 – Curva de Lorenz do volume de depósito das Mesorregiões – Região Norte e demais UFs.....	155
Figura A - 3.2 – Curva de Lorenz do volume de crédito das Mesorregiões – Região Norte e demais UFs.....	156
Figura A - 3.3 – Curva de Lorenz de agências das Mesorregiões – Região Norte e demais UFs	157
Figura A - 3.4 – Submercados de Campo Grande - MS	160
Figura A - 3.5 – Mapas de submercado – Municípios	161
Figura A - 3.6 – Mapas de submercado – Municípios da região NE.....	162
Figura A - 3.7 – Mapas de submercado – Municípios da região SE.....	162
Figura A - 3.8 – Mapas de submercado – Municípios da região S	163
Figura A - 3.9 – Histograma da participação dos bancos no mercado de Crédito segundo a cobertura geográfica (Nacional ou Local)	164
Figura A - 3.10 – Histograma da participação dos bancos no mercado de Agências segundo a cobertura geográfica (Nacional ou Local)	165

Apresentação

Nosso objetivo é aprofundar os conhecimentos acerca da economia bancária através de estudos empíricos aplicados ao Brasil. Organizamos este trabalho em três capítulos, sem contar esta breve introdução. Cada um é independente e trata de tópico distinto, compreendendo um estudo com introdução, desenvolvimento, conclusão, referências bibliográficas e apêndice. Em comum, todos eles tratam de microeconomia bancária e utilizam dados desagregados provenientes do Banco Central do Brasil. Os dois primeiros referem-se ao mercado de crédito e o terceiro testa as predições da teoria dos submercados independentes (Sutton, 1998).

As variáveis que afetam a demanda por crédito constituem o tema do primeiro capítulo. Motivado pelo fato de que as taxas de juros cobradas nas operações de crédito no Brasil são bastante elevadas, o objetivo desse trabalho é verificar a resposta do consumidor às variações nas taxas de juros e no prazo de pagamento, quando opta por contratar um financiamento. Se os consumidores sofrem restrição ao crédito então deveriam responder mais à segunda e menos à primeira variável. A parte empírica do trabalho utiliza uma rica fonte de dados do Banco Central do Brasil: o Sistema de Informações de Crédito – SCR. Em particular, utilizamos os dados das operações iniciadas em mai/2004, na modalidade Financiamento de Veículos para pessoas físicas. Encontramos evidências de que os consumidores respondem tanto a variações nas taxas de juros quanto no prazo de pagamento. Dessa forma, para que o financiamento seja contratado, além da taxa de juros é importante que o valor da prestação seja compatível com o orçamento do consumidor. Este resultado se repete quando dividimos a amostra em grupos mais e menos prováveis de se defrontarem com a restrição.

O segundo estudo tem como objetivo analisar a relação entre taxa de juros e garantias bem como a influência do tempo de relacionamento nos contratos de crédito. Existem diversas razões para que a taxa de juros e a garantia sejam negativamente relacionadas. Tudo o mais constante, parece intuitivo que um empréstimo com garantia seja menos arriscado que um sem. Na presença de assimetria informacional, a garantia é uma importante ferramenta de sinalização, uma vez que os empresários que escolhem projetos de menor risco são mais propensos a escolher um empréstimo com garantia em troca de uma taxa de juros mais baixa.

A garantia também ajuda a mitigar os efeitos do risco moral depois que o crédito é concedido. No entanto, esses resultados não são necessariamente compartilhados pela literatura empírica (Berger e Udell, 1990; Jiménez e Saurina, 2004). O maior tempo de relacionamento entre o cliente e o banco, por sua vez, pode ajudar a mitigar os problemas da assimetria informacional, reduzindo os valores cobrados da taxa de juros e da garantia. Para estimar as regressões utilizamos novamente os dados do SCR, desta vez para as operações iniciadas em set/2005 na modalidade Capital de Giro. Diversas técnicas econométricas são empregadas. Os resultados empíricos indicam que a taxa de juros e as garantias são negativamente relacionadas nos contratos de crédito e o um tempo de relacionamento mais longo ajuda a diminuir os problemas decorrentes da assimetria informacional.

Por fim, no terceiro estudo pretendemos contribuir para a literatura de estrutura de mercado bancário brasileira. Testamos as predições do modelo de submercados independentes de Sutton (1998) utilizando os dados das agências bancárias provenientes do ESTBAN do Banco Central do Brasil (dez/2002). Nesta teoria, o equilíbrio do agregado de submercados ocorre com um grau mínimo de desigualdade no tamanho das firmas. Por isso, os dados do mercado bancário para o Brasil ou para as UFs podem revelar grande assimetria no tamanho das firmas que é decorrente da agregação de diversos submercados independentes. Inicialmente, o esboço das curvas de Lorenz é desenhado para diversas desagregações territoriais. Depois, tentamos identificar qual seria a melhor extensão geográfica que representaria um submercado (microrregiões ou municípios), bem como distinguir os submercados em municípios compostos por mais de um deles. Os resultados para o mercado de agências bancárias são consistentes com as predições da teoria. O Brasil, as UFs e as mesorregiões apresentam grande desigualdade no tamanho das firmas. Por sua vez, as microrregiões e os municípios parecem representar submercados independentes.

Nota sobre tabelas, figuras, equações e apêndices.

Este trabalho está dividido em três capítulos (Capítulos 1, 2 e 3), excluindo esta apresentação. Os objetos (tabelas e figuras) são identificados através de números, adotando um sistema de dois dígitos, separados por um ponto. O primeiro dígito refere-se ao número do Capítulo e o segundo, ao número do objeto. Dessa forma, a primeira tabela do Capítulo 1 é identificada

como **Tabela 1.1**, a primeira figura do Capítulo 2 é identificada como **Figura 2.1** e assim por diante.

No Apêndice, a identificação dos objetos foi feita de maneira análoga acrescentando-se a letra “A”. A primeira tabela do Apêndice do Capítulo 1 é identificada como **Tabela A – 1.1**, a primeira figura do Apêndice do Capítulo 3 é identificada como **Figura A – 3.1** e assim sucessivamente.

Para que o texto não ficasse muito carregado, somente as principais equações e as que foram referenciadas no corpo do trabalho foram identificadas, seguindo o mesmo critério das tabelas e das figuras.

De maneira similar, os Apêndices são identificados pelo número do capítulo seguido de um ponto e uma letra do alfabeto. Assim, o **APÊNDICE 1.A** refere-se ao Apêndice A do Capítulo 1. Para facilitar a localização dos Apêndices, o número da sua página foi incluído no momento que o mesmo é mencionado no texto.

Capítulo 1

1 Mercado de Crédito no Brasil: um Estudo sobre os Impactos da Taxa de Juros e do Prazo de Pagamento nos Financiamento de Veículos para Pessoas Físicas

1.1 INTRODUÇÃO

Este trabalho tem como objetivo testar empiricamente se os consumidores sofrem restrição de crédito no mercado de financiamento de veículos brasileiro. A motivação para este estudo advém do fato de que, no Brasil, existe a percepção de que as pessoas se preocupam apenas se o valor da prestação é compatível com o orçamento mensal e, portanto, se importam menos com as taxas de juros embutidas nas prestações. Se o consumidor se comporta desta maneira, então uma possível explicação é que o mesmo sofre restrição de crédito.

No mercado de crédito, os contratos podem ser caracterizados pelas variáveis: valor do empréstimo, taxa de juros, prazo de pagamento e garantia. Tudo o mais constante, um consumidor aumenta a sua quantidade emprestada quando houver uma redução na taxa de juros e/ou um aumento no prazo de pagamento. Se a demanda por empréstimo é menos sensível à taxa de juros que ao prazo de pagamento, uma possível interpretação é que a importância deste é superior à da taxa de juros na definição do valor da prestação. No entanto, no Brasil as taxas de juros são bastante elevada e, por isso, é possível que a sua variação afete o comportamento dos consumidores, contrariando o senso comum.

Sob o ponto de vista das instituições financeiras, quanto maior o prazo de pagamento, maior tende a ser a incerteza no recebimento e, portanto, maior a taxa de juros cobrada. Ainda que o próprio veículo financiado sirva de garantia no contrato de empréstimo, a preocupação dos

credores refere-se ao percentual de comprometimento da renda mensal com o pagamento das prestações. A probabilidade de inadimplência aumenta quanto maior a destinação da renda mensal e o número de meses em que o consumidor irá pagar o financiamento. Por isso, o prazo de pagamento é também uma importante variável na definição do limite de crédito.

Attanasio, Goldberg e Kyriazidou (2008), AGK daqui em diante, argumentam que a demanda por empréstimos dos indivíduos que não se defrontam com a restrição ao crédito deve ser função da taxa de juros (preço), mas independente do prazo de pagamento. Por outro lado, os consumidores que têm restrição de crédito devem ser menos sensíveis a variações nos preços e mais sensíveis a variações no prazo de pagamento.

Para testar empiricamente a presença de restrição de crédito, Jappelli *et al.* (1998) propõem uma estimação em dois estágios utilizando bases de dados distintas: 1) inicialmente estima-se a probabilidade de o consumidor sofrer a restrição utilizando o *Survey of Consumer Finances*; e 2) essa probabilidade estimada é utilizada na equação do segundo estágio, que utiliza os dados do *Panel Study of Income Dynamics*. Utilizando microdados de cartão de crédito, Gross e Souleles (2001) testam qual a variação da dívida do consumidor com relação a um aumento na sua linha de crédito, cujo coeficiente pode ser interpretado como a propensão marginal do consumo sobre o aumento da liquidez. Jappelli (1990), por sua vez, utiliza dados onde os consumidores solicitaram crédito bancário, mas tiveram ou poderiam ter os seus pedidos negados pelas instituições financeiras.

Como em geral os indivíduos que sofrem restrição não são diretamente observados, diversos autores inferem sobre o seu comportamento dividindo a amostra em duas partes, uma contendo os indivíduos que são mais prováveis e outra com os menos prováveis de sofrer a restrição (Bernanke, 1984; Hayashi, 1985; Zeldes, 1989; AGK, 2008). Esta será a abordagem utilizada em nossas estimações.

Para atender ao objetivo proposto, dispomos de uma *cross-section* dos financiamentos de veículos para pessoas físicas do SCR – **Sistema de Informações de Crédito do Banco Central do Brasil**. O SCR é um cadastro de informações positivas gerenciado pelo Bacen, cuja base de dados é alimentada mensalmente pelas instituições financeiras (IFs). Esta base inclui dados referentes a operações cuja responsabilidade total é maior ou igual a R\$ 5.000,00, a vencer e vencidas, bem como os valores referentes às fianças e aos avais prestados pelas

instituições financeiras a seus clientes. Esses dados são compartilhados com as IFs participantes, contribuindo para uma melhor gestão de risco do crédito. Em particular, obtivemos os dados de financiamento de veículos para pessoas físicas – PF iniciados em maio/2004. Para o estudo proposto, esta modalidade se mostra adequada devido aos seguintes fatores: 1) por ser um bem relativamente caro, é provável que boa parte das aquisições de veículos seja financiada; 2) o valor das parcelas compromete parte significativa da renda e por isso os termos do contrato *ex ante* tenderiam a ser mais bem avaliados pelos compradores; e 3) a princípio (salvo diferenças nos valores oferecidos de entrada), todos os tomadores possuem garantia, que é o próprio veículo financiado. Por isso, as mesmas não deveriam influir significativamente na taxa de juros e no prazo do empréstimo contratado.

Apesar de o SCR ser uma base de dados bastante rica, ele não traz o percentual do veículo financiado, nem as informações demográficas e socioeconômicas das pessoas que adquiriram o financiamento.

A abordagem proposta neste trabalho segue AGK (2008): será estimada uma equação de demanda por empréstimos e, de acordo com a disponibilidade dos dados, iremos considerar as peculiaridades dos contratos de financiamento de automóveis. Neste mercado, é possível que exista um viés de seleção na amostra, pois os contratos de financiamento estão disponíveis somente para as pessoas que compraram e financiaram a aquisição. Adicionalmente, tanto a taxa de juros quanto o prazo de pagamento são potencialmente endógenos nos contratos. Para a estimação, será utilizada a metodologia semiparamétrica proposta por Das, Newey e Vella (2003).

Os resultados empíricos não sustentam o fato estilizado de que os consumidores dão pouco valor à taxa de juros embutidas nas prestações. O comprador responde tanto a juros quanto ao prazo de pagamento, cujas elasticidades calculadas são -0,45 e +0,33, respectivamente. Para encontrar evidência da presença da restrição de crédito, dividimos a amostra em grupos menos e mais prováveis de se defrontar com este problema. Um padrão parece emergir: os grupos mais prováveis apresentam sensibilidade maior tanto à taxa de juros quanto ao prazo de pagamento que os menos prováveis. O resultado da elasticidade do prazo de pagamento está de acordo com o esperado pela teoria. A maior elasticidade da taxa de juros neste caso, comparada com o resultado obtido por AGK (2008), pode ser explicada pelo fato de que, no Brasil, as taxas são bastante elevadas e por isso uma pequena variação pode gerar valores

significativamente diferentes nas prestações. Os resultados obtidos não são contraditórios com a presença de restrição ao crédito. Conclui-se então que existem evidências de que os consumidores sofrem restrição ao financiamento de veículos. No entanto, respondem não somente a variações no prazo de pagamento, mas também à taxa de juros.

1.2 BREVE DESCRIÇÃO DO MERCADO DE FINANCIAMENTO DE VEÍCULOS NO BRASIL

Esta seção tem como objetivo apresentar alguns dos principais dados referentes ao financiamento de veículos no Brasil e que corroboram a importância de uma melhor compreensão deste mercado.

O número de emplacamentos de veículos tem aumentado bastante nos últimos anos. Entre 2002 a 2006, o número de automóveis emplacados no ano aumentou de 1,18 milhões para 1,56 milhões e o de comerciais leves aumentou de 208 mil para 274 mil. Nos oito primeiros meses de 2007, foram emplacados 1,24 milhões de automóveis e 220 mil veículos do tipo comercial leve, praticamente os mesmos números de todo o ano de 2004 (**Tabela 1.1**).

Tabela 1.1 – Emplacamento de automóveis e comerciais leves

Ano	Automóveis	Com. Leves
2002	1.182.317	207.827
2003	1.172.901	178.088
2004	1.258.821	220.478
2005	1.369.093	251.236
2006	1.556.231	274.171
2007*	1.237.733	220.338

* De janeiro a agosto.

Fontes: 2002 a 2005 – Denatran *in*: Anuário Fenabrave 2005.

2006 e 2007 – Renavan *in*: Informativos Fenabrave nº108 e nº124.

Dentre as operações de crédito para Pessoa Física, as mais importantes são o Crédito Pessoal e Aquisição de Veículos, cujo volume vem crescendo acentuadamente, principalmente a partir de 2004 (**Figura 1.1**).

Figura 1.1– Volume de operações de crédito – Pessoa Física – R\$ milhões

Fonte: **Bacen**.

Conforme dados do Banco Central do Brasil - Bacen, verifica-se que o Cheque Especial com juros entre 140 a 180% a.a. é a modalidade que apresenta maiores taxas entre as operação com taxas pré-fixadas para Pessoas Física. Por outro lado, a Aquisição de Veículos apresenta a taxa mais baixa, mas ainda assim cerca de duas vezes maior do que a taxa SELIC (**Figura 1.2**). Uma das razões para a menor taxa de juros cobrada é que, em geral, nos financiamentos o próprio veículo é dado como garantia, e fica alienado ao banco até o término do contrato. Uma pesquisa realizada na *internet*¹ mostra que além da alienação do bem, a nota promissória ou aval também são aceitos como garantia. Na maioria dos casos exige-se entrada de no mínimo 20 a 30% do valor do veículo financiado. Para alguns bancos, a liberação do financiamento está atrelada à condição de que o contratante seja correntista da instituição. Adicionalmente, o veículo financiado deve ter idade máxima de 7 a 12 anos. Esses fatores contribuem para que, em geral, as taxas de juros para financiamento de automóveis sejam menores do que em outras modalidades de crédito.

¹ Os seguintes sites de instituições financeiras foram consultados em 10/10/2007: www.bradesco.com.br; www.italu.com.br, www.hsbc.com.br, www.bancoreal.com.br, www.bb.com.br e www.santander.com.br.

Figura 1.2 – Taxa de juros pré-fixadas – PF e Selic acumulada no mês (%a.a.)

Fonte: **Bacen**.

Nos últimos anos, o refinanciamento e a “troca com troco” têm sido recorrentemente utilizados para levantar recursos para quitar outras dívidas, dado o fato de a taxa de juros de financiamento de veículos ser inferior às das demais modalidades de crédito. No refinanciamento, um automóvel preferencialmente já quitado é alienado à financeira. Na “troca com troco”, na compra de um modelo mais novo, em vez de dar o carro usado de entrada, vende-se o mesmo à loja e financia-se 100% do carro adquirido.

Conforme observado na **Figura 1.3**, dentre as formas de pagamento, o financiamento é a mais utilizada, seguido pelo pagamento à vista. A participação do *leasing* cresceu bastante nos últimos anos, passando de 3% em 2001 para 24% no 1ºsem/2007. A forma menos utilizada proporcionalmente é o consórcio, cuja participação encolheu de 19% em 2000 para 4% no 1ºsem/2007.

Figura 1.3 – Forma de pagamento na venda de veículos

Fonte: Anef.

A média mensal de novas concessões de crédito para Aquisição de Veículos – PF com juros pré-fixados apresenta tendência crescente, conforme os dados da **Tabela 1.2**. Esse valor passou de R\$ 4.349 milhões em 2006 para R\$ 5.119 milhões no 1º semestre de 2007. O prazo médio das operações de financiamento de veículos – PF em carteira também tem se dilatado e no 1º semestre de 2007 foi de 18,4 meses, o mais longo dos últimos anos. Entre os anos 2001 e 2007, o Plano Máximo de empréstimo (os planos de financiamento disponibilizados ao consumidor) aumentou de 48 para 72 meses. Acompanhando o alargamento do Plano Máximo, o Plano Médio (os planos utilizados), também aumentou de 34 para 41 meses.

Tabela 1.2 – Aquisição de veículos

Ano	Novas concessões – média mensal (R\$ milhões)	Prazo Médio (meses)	Plano Máximo (meses)	Plano Médio (meses)
2001	1.776	17,5	48	34
2002	1.679	17,1	48	34
2003	2.045	16,9	48	34
2004	2.759	15,5	60	34
2005	3.498	15,8	60	36
2006	4.349	17,1	72	39
1ºsem/2007	5.119	18,4	72	41

Prazo Médio: Reflete as operações em carteira da modalidade Aq. de bens PF – veículos.

Plano Máximo: São os planos de financiamento disponibilizados pelo mercado ao consumidor.

Plano Médio: É a ponderação dos planos de financiamento utilizados pelo consumidor.

Fontes: Anef (Plano máximo e Plano médio) e Bacen.

Pelos dados da **Tabela 1.3**, verifica-se que os atrasos no pagamento do financiamento para Aquisição de Veículos – PF sofrem pequenas oscilações ao longo do tempo. Aproximadamente 90% é pago sem atraso, entre 3-4% tem atraso de 15-30 dias e cerca de 3% apresenta atraso superior a 90 dias.

Tabela 1.3 – Faixas de atraso – Aquisição de veículos – Pessoa Física – R\$ milhões

Data	Sem atraso		15 a 30 dias		31 a 90 dias		> 90 dias		Saldo Total
	Saldo	%	Saldo	%	Saldo	%	Saldo	%	
dez/00	14.421	92,5%	465	3,0%	428	2,8%	283	1,8%	15.597
dez/01	22.056	89,3%	930	3,8%	975	4,0%	735	3,0%	24.697
dez/02	24.102	89,5%	1.049	3,9%	997	3,7%	776	2,9%	26.924
dez/03	26.818	89,4%	1.142	3,8%	1.066	3,6%	959	3,2%	29.984
dez/04	35.184	92,4%	1.124	3,0%	954	2,5%	801	2,1%	38.063
dez/05	46.289	91,4%	1.614	3,2%	1.654	3,3%	1.115	2,2%	50.671
dez/06	57.056	89,9%	2.084	3,3%	2.256	3,6%	2.077	3,3%	63.472
ago/07	66.776	90,0%	2.438	3,3%	2.617	3,5%	2.366	3,2%	74.198

Fonte: Bacen.

1.3 ARCABOUÇO TEÓRICO

As três principais variáveis que compõem um contrato de financiamento de veículos são: o valor do empréstimo, a taxa de juros e o prazo de pagamento. Tudo o mais constante, quanto maior a taxa de juros, menor o valor que os consumidores desejariam tomar emprestado e quanto maior o prazo de pagamento, maior esse valor. Se, pelo lado dos consumidores, um prazo de pagamento longo implica valores de prestação menores, do ponto de vista dos bancos, quanto maior o prazo de financiamento, maior a incerteza e, portanto, maior tende a ser a taxa de juros cobrada.

Em geral, os financiamentos são concedidos apenas para os veículos novos e semi-novos e o próprio bem serve de garantia². Como todos os contratos tendem a apresentar a mesma garantia, sob a ótica do prestador, a capacidade de pagamento dos tomadores é um dos fatores mais importantes para a concessão do crédito. Os bancos somente estão dispostos a financiar o consumo presente para as famílias cujo valor da prestação esteja de acordo com seus rendimentos.

Conforme os dados apresentados na **Seção 1.2**, cerca de 40-50% dos veículos são financiados e cerca de 30% são pagos à vista. Apesar de a taxa de financiamento de um veículo ser inferior a de outras modalidades de crédito, ainda assim o seu valor é em média o dobro da taxa SELIC. Um indício de que nesse mercado o consumidor sofre restrição ao crédito é obtido se a resposta a variações no prazo de pagamento é maior que a variações na taxa de juros, conforme argumento de AGK (2008).

Estes autores utilizam duas definições para o termo “restrição ao crédito”. A primeira, com sentido mais estrito, diz que um consumidor sofre restrição ao crédito se ele não pode tomar emprestada a quantia que gostaria de modo a financiar o consumo presente utilizando recursos que poderia auferir no futuro. A segunda definição afirma que um consumidor se defronta com restrição ao crédito se a taxa de juros a qual toma emprestado é superior à taxa a qual aplica os recursos. Nesse trabalho, adotamos essas mesmas definições.

² Ainda que as garantias diminuam as perdas dos bancos em caso de inadimplência, a sua execução depende do tempo e do custo de sua recuperação. Nesse trabalho não consideramos os possíveis efeitos do judiciário.

Um automóvel é um bem durável relativamente caro e por isso em geral é adquirido através de acúmulo de poupança ou contratação de financiamento. Indivíduos com renda muito baixa estão excluídos do mercado de automóveis novos e semi-novos e, portanto, excluídos do mercado de financiamento. E, se uma pessoa não tem o crédito aprovado porque apresenta uma renda incompatível com o bem financiado, não significa que a mesma sofra restrição ao crédito.

Um critério objetivo para avaliar a capacidade de pagamento é calcular o percentual de comprometimento da renda mensal com o pagamento das parcelas. Se os bancos o utilizam para negar crédito, então, uma variação nos termos do contrato (taxa de juros ou prazo) pode fazer com que um consumidor deixe de (ou passe a) sofrer restrições. Tudo o mais constante, uma redução na taxa de juros diminui o valor da prestação. De maneira análoga, um prazo de pagamento mais extenso implica valores de prestações individuais menores. No entanto, o prazo máximo de financiamento é limitado pela depreciação física do veículo dado como garantia e os consumidores devem quitar o bem em prazo inferior à sua depreciação, pagando antecipadamente por um serviço que irá efetivamente receber no futuro³.

Dessa forma, se uma variação na taxa de juros ou no prazo de pagamento alterar os limites de crédito do consumidor, então é possível que o mesmo responda a ambas variáveis na definição do valor contratado, mesmo que se defronte com restrição.

1.3.1 Modelo Simples

Conforme **AGK** (2008), o estudo da demanda por automóveis na presença de restrição ao crédito requer um modelo dinâmico no qual os consumidores alocam seus recursos ao longo do tempo. No entanto os modelos padrão de ciclo de vida que incorporam incerteza e ativos e passivos com diferentes taxas de juros somente possuem solução fechada para a demanda por empréstimos sob certas circunstâncias. Por isso os autores optam por extrair algumas intuições do comportamento dos consumidores na presença de restrição ao crédito das condições de primeira ordem do modelo de ciclo de vida padrão.

³ Por isso a decisão de financiamento afeta as alocações intertemporais ótimas de consumo. No entanto, para a parte empírica dispomos de dados de uma *cross-section* (mai/2004) e por isso optamos pelo uso da estática comparativa.

No nosso caso, para a parte empírica dispomos de dados de uma *cross-section* e por isso adotaremos uma abordagem estática para inferir sobre o comportamento teórico dos consumidores. Nosso pressuposto é que o consumidor que opta pelo financiamento sabe quais são a sua renda e a sua despesa e, portanto, quanto lhe sobra mensalmente para pagar as prestações. E, dado que não espera grandes oscilações na renda, esse orçamento deve se repetir durante todo o período de pagamento.

As decisões dos consumidores podem ser modeladas como uma variável binária d , tal que $d = 1$ se compra e financia e $d = 0$ se compra, mas não financia. Para o consumidor que deseja comprar e financiar define-se $U(.)$ a utilidade que depende de um vetor C de consumo de outros bens e dos gastos $g(G)$ com combustível, impostos, etc. G é o valor do automóvel, T ⁴ é o prazo de pagamento do financiamento, $\phi \in (0,1]$ a proporção do veículo financiado e $P(\phi, G, r, T)$ é o valor da prestação. Se o consumidor sofre restrição ao crédito, um alongamento no prazo de pagamento permite diminuir o valor de P .

Se o consumidor comprar e financiar $d = 1$ e $\phi \cdot G > 0$, e se comprar, mas não financiar $d = 0$ e $\phi \cdot G = 0$. Se não compra o automóvel, $g(G = 0) = 0$.

Quando o consumidor não adquire o veículo financiado ($\phi \cdot G = 0$), ele pode gastar a totalidade da sua renda (M) com o bem C . Se optar por adquirir o automóvel ($\phi \cdot G > 0$) deve dividir a renda com o pagamento da prestação e com os bens C e g .

Do ponto de vista do credor, dado que a princípio todos os contratos de financiamento apresentam garantia, a capacidade de pagamento é a principal variável na concessão do crédito⁵. Dado que os juros e o prazo de pagamento são prefixados, o valor de cada prestação

⁴ Para que o veículo sirva de garantia, T deve ser inferior ao tempo de depreciação do veículo.

⁵ Em geral, nos empréstimos para pessoas jurídicas, os modelos teóricos baseiam-se em um jogo, onde os bancos financiam os projetos dos empresários que podem ser do tipo bom ou ruim e que podem investir em projetos arriscados ou seguros, com diferentes retornos. Neste caso, a capacidade de pagamento da dívida está atrelada à escolha do projeto (arriscado ou seguro) pelo empresário. No caso de empréstimos para pessoa física, por sua vez, a capacidade de pagamento está atrelada ao comprometimento mensal da renda no pagamento da dívida, dado que se espera que a mesma não sofra variações bruscas de um mês para o outro. E o melhor previsor para a renda futura é a renda atual, dado que os gastos são fixos.

é fixo. Nesse caso, a prestação mensal que não onera o orçamento é tal que: $P(\phi, G, r, T) \leq M - p_c C - p_g G$, onde p_c é o preço do bem C , p_g o preço do bem g e M é a renda. No momento em que se concede o crédito, supõe-se que essa restrição orçamentária se repetirá durante os T períodos de quitação do financiamento.

O problema do consumidor é encontrar um contrato em que $P(\phi, G, r, T)$ seja compatível com a sua restrição orçamentária e garanta um nível mínimo de utilidade. Seja $U^{\min}$ o nível de utilidade mínima aceito pelo consumidor durante o pagamento das prestações tal que $U^{\min}(C, g / \phi, G > 0) \geq U(C / G = 0)$. Dados os preços, o consumidor escolhe a despesa mínima com os bens C e g que atinja esse nível de utilidade (função gasto).

O problema do consumidor pode ser escrito como:

$$\begin{aligned} E(p_c, p_g, U / \phi, G > 0) &= \text{Min}_{C, g} p_c C + p_g g \\ \text{s.a. } U(C, g / \phi, G > 0) &\geq U^{\min} \end{aligned}$$

Onde p_c e p_g são os preços do bem C e g , respectivamente.

A solução desse problema é a demanda Hicksiana $C^h(p_c, p_g, U / \phi, G > 0)$ e $g^h(p_c, p_g, U / \phi, G > 0)$, cuja despesa mínima é dada por: $B^{\min} = p_c C^h + p_g g^h \Rightarrow B^{\min} = p_c C^h(p_c, p_g, U / \phi, G > 0) + p_g g^h(p_c, p_g, U / \phi, G > 0)$.

Por simplificação, considere que a despesa mínima é um valor fixo, $B^{\min} = \bar{B}$ e é uma proporção $1 - \gamma$ da renda, onde $0 < \gamma < 1$ é a fração que pode ser gasta com a prestação. Temos então que:

Equação 1.1: $P_{\max}(\phi, G, r, T) \leq \gamma M$.

Na **Figura 1.4**, representamos dois valores de parcela, um mais alto $\left(\bar{\gamma} M\right)$, representado pelo trapézio vermelho e outro mais baixo $\left(\gamma_- M\right)$, representado pelo trapézio azul, e as suas respectivas curvas de utilidade. A restrição orçamentária antes do pagamento da parcela é representada pela linha preta.

Figura 1.4: Restrição orçamentária do consumidor que compra e financia.

Seja $\bar{\gamma}$ a fração máxima de comprometimento da renda que o banco considera aceitável com a prestação. Consumidores com $\gamma \cong \bar{\gamma}$ são os candidatos a defrontarem-se com a restrição ao crédito. Como a renda e $\bar{\gamma}$ são exógenos, uma maneira de esses consumidores terem acesso ao crédito é alterando o lado esquerdo da **Equação 1.1**⁶. Se tanto a taxa de juros quanto o prazo influenciarem o valor do limite de crédito, então uma pequena variação (negativa) na taxa de juros e/ou (positiva) no prazo de pagamento pode fazer com que uma pessoa passe a ter acesso ao crédito.

Da **Equação 1.1** temos que: $\phi.G = h(r, T, \gamma M; \beta)$ onde $\phi.G$ é o valor do empréstimo, r é a taxa de juros, T é o prazo de pagamento e β são os parâmetros.

Defina $P(\phi.G, r, T) = \phi.G \left[\frac{(1+r)^T \cdot r}{(1+r)^T - 1} \right]$, temos então:

⁶ Suponha um consumidor que possui renda de R\$ 1 mil e deseje financiar um veículo de R\$ 10 mil. Dada as suas características, o banco avalia que o comprometimento mensal máximo da renda com prestações deve ser de R\$ 300. Um contrato com taxa de juros de 2% a.m. e prazo de 48 meses gera uma prestação de R\$ 326. Neste caso, uma pequena diminuição na taxa de juros e/ou alongamento do prazo de pagamento fazem com que o cliente venha a ter o crédito aprovado.

$$\phi.G \left[\frac{(1+r)^T \cdot r}{(1+r)^T - 1} \right] = \gamma M \Rightarrow \phi.G = \gamma M \left[\frac{(1+r)^T - 1}{(1+r)^T \cdot r} \right]$$

E as derivadas parciais são: $\frac{\partial \phi.G}{\partial M} > 0$; $\frac{\partial \phi.G}{\partial T} > 0$ e $\frac{\partial \phi.G}{\partial r} < 0$.

Segue então que tudo o mais constante, quanto maior o prazo de pagamento maior o volume de empréstimo; quanto menor a taxa de juros, maior o volume de empréstimo; e quanto maior a renda, maior o volume de empréstimo.

Espera-se que para um consumidor que sofre restrição ao crédito, a magnitude da resposta com relação a variação na taxa de juros ou no prazo de pagamento deve ser maior que outro que não se defronta com a restrição. Isso porque para o grupo dos “restritos”, uma diminuição na taxa de juros ou um aumento no prazo de pagamento pode representar a diferença entre ter ou não crédito. Por isso, o fato de os consumidores responderem a variações na taxa de juros não implica necessariamente que neste mercado não está presente a restrição de crédito.

Para a parte empírica do trabalho, o ideal seria ter os dados das pessoas que não tiveram o crédito aprovado. Como não dispomos desses dados, utilizamos uma abordagem frequente na literatura e dividimos a amostra em grupos mais ou menos prováveis de ter acesso ao crédito, utilizando as informações provenientes do SCR dos empréstimos iniciados em maio/2004.

1.4 MODELO EMPÍRICO

No final da **subseção 1.3.1** derivamos os sinais esperados das principais variáveis de um contrato de empréstimo. A equação de demanda que queremos estimar pode ser representada da seguinte forma:

Equação 1.2: $\phi G = \gamma x + f(r, T) + \varepsilon$

Onde ϕG é o valor que o consumidor deseja financiar, r é o valor da taxa de juros, T é o prazo do financiamento, x é o vetor de variáveis exógenas e ε é o termo de erro não observado.

A estimação da **Equação 1.2** requer tratamento acerca da seleção da amostra e da endogeneidade da taxa de juros e do prazo de pagamento, que serão discutidos a seguir.

1.4.1 Seleção da amostra

Para estimar a **Equação 1.2**, deve-se considerar o problema de seleção amostral, pois o valor do empréstimo é observado somente para os consumidores que financiaram a aquisição. A primeira decisão do consumidor a ser modelada deve ser se “compra” ou “não compra” o automóvel, seguida pela decisão de “financiar” ou “não financiar” a aquisição. Neste caso, não encontramos nenhuma variável que afete a decisão de comprar, mas não de financiar; e, seguindo o procedimento adotado por AGK (2008) colapsamos o problema de seleção em uma única decisão onde a variável dependente é igual a 1 se “compra e financia” e zero, caso contrário.

Apesar de ser uma fonte rica de dados dos contratos, a base do SCR não traz os dados socioeconômicos nem demográficos das pessoas que compraram e financiaram a aquisição do veículo. Para tratar da seleção amostral, utilizamos a base de dados da Pesquisa de Orçamentos Familiares 2002-2003 (POF 2002-2003) para imputar a probabilidade de comprar e financiar um automóvel na base do SCR. A única variável disponível para fazer a ligação

entre as duas bases é a região⁷ da agência bancária que concedeu o empréstimo (proveniente do SCR) e a região de domicílio das famílias (POF 2002-2003)⁸. A POF traz as *Despesas Individuais com Aquisição de Veículos no Período de Referência de 12 Meses*. Como estamos interessados em modelar apenas o financiamento do automóvel, e não a decisão da sua compra, consideramos que um possível candidato a mercado das famílias que potencialmente poderiam optar por financiar a compra do automóvel (mai/2004) é composto pelas famílias que adquiriram o automóvel nesse período de referência de 12 meses da POF. Utilizamos os dados dos indivíduos que declararam ter adquirido moto e automóvel (passeio e utilitário, nacional e importado) e definimos a variável $d = 1$ se o indivíduo comprou e financiou (pagamento a prazo para a Unidade de Consumo (UC) ou para Outra UC) e $d = 0$ se comprou, mas não financiou (pagamento monetário à vista para a UC ou para Outra UC).

1.4.2 Variáveis endógenas

Num contrato de financiamento de automóveis é provável que o valor do empréstimo, a taxa de juros e o prazo de pagamento sejam determinados conjuntamente. Contratos que apresentam maiores prazos e valores de empréstimo devem ter maiores taxa de juros. Os pacotes de financiamentos ofertados dependem do valor da prestação que o cliente pode arcar mensalmente, devendo ser correlacionado com as suas características. Por isso, na **Equação 1.2**, a taxa de juros e o prazo de pagamento devem ser determinados de maneira endógena.

Um experimento natural ou um fato não intencional que produz uma variação exógena na taxa de juros ou no prazo de pagamento é um candidato a variável instrumental. Para instrumentalizar a taxa de juros, AGK (2008) utilizam a mudança na legislação do imposto de renda, que extinguiu gradativamente no período de cinco anos a dedutibilidade dos juros dos financiamentos. Como instrumento para o prazo de pagamento, os autores utilizam o crescente aumento da durabilidade dos veículos, calculado para cada modelo utilizando os dados do *Consumer Expenditure Survey* (CEX). Se o próprio veículo serve de garantia, à medida que a sua durabilidade aumenta, o prazo de pagamento pode ser alongado.

⁷ Dividimos cada uma das 27 UF's em capital e resto, totalizando 54 regiões.

⁸ Nesse caso adotamos a hipótese de que o consumidor irá escolher uma agência bancária para fazer o financiamento na região onde reside.

Utilizando dados de um experimento no qual os clientes de uma instituição de microcrédito sul-africana receberam correspondência com valores de taxa de juros e do prazo de pagamento aleatórios (dado o risco) e valor do empréstimo igual ao último valor contratado, Karlan e Zinman (2008) estimam a resposta do consumidor a variações nas taxas de juros e no prazo de pagamento dos empréstimos. Os autores utilizam como instrumento para o prazo de pagamento os valores sugeridos nas correspondências.

Em geral, para se estimar uma equação de demanda em que as quantidades são função dos preços, pressupõe-se que todos os consumidores defrontam-se com o mesmo preço. Se isso não ocorre, o uso de preços médios praticados no mercado ocasiona erros de medida⁹, razão pela qual deve haver correlação entre a variável explicativa e o termo de erro. Na base de dados do SCR, observamos os valores dos empréstimos, da taxa de juros e do prazo de pagamento para cada transação e por isso, se considerarmos que os consumidores são tomadores de preços, esta forma de endogeneidade é menos preocupante na estimação.

Outra forma de endogeneidade pode advir do fato de que características individuais não observadas estejam correlacionadas com a taxa de juros e o prazo¹⁰ (viés de variável omitida). Conseguimos controlar parcialmente a heterogeneidade não observada, pois a base do SCR apresenta características individuais tais como tempo de relacionamento e número de IFs com a qual o tomador está relacionado.

Nossa fonte de instrumentos para os coeficientes da equação de demanda será a equação de oferta. Suponha que os bancos travem uma concorrência de preços do tipo *Bertrand*. Se existem $j = 1, \dots, J$ bancos e $t = 1, \dots, T$ mercados, a condição de primeira ordem para o preço é dada por (vide Nakane *et al.*, 2006):

$$p_j^e = r + c_j^E - \frac{E_j}{V^E \sum_{t=1}^T M_t \frac{\partial s_{jt}^E}{\partial p_j^E}}$$

⁹ Vide Nevo, 2001 e 2000.

¹⁰ Agradecemos ao prof. Gabriel Madeira por nos informar acerca deste fato.

onde:

p_j^e é a taxa de juros de empréstimo do banco j ;

r é a taxa de juros dos títulos do banco j ;

c_j^E é o custo marginal do banco j de prover uma unidade de empréstimo;

E_j é o volume de empréstimos concedidos pelo banco j ;

V^E é o valor médio das operações de empréstimo;

M_t é o tamanho do mercado t ;

s_{jt}^E é o *market share* em termos de empréstimos do banco j no mercado t .

Qualquer variável exógena que desloca o lado esquerdo da curva de oferta é um candidato à variável instrumental. Dessa forma, utilizamos como instrumento variáveis que indicam o grau de competição e o potencial do mercado, sendo que estas variáveis estão definidas por município. São elas: o índice de *Herfindahl* para empréstimos (*hempr*), a razão crédito/PIB (*pot*) e o número de agências por 1000 habitantes (*nagencph*). Como não é possível identificar a IF, os possíveis deslocadores de oferta (variáveis de custo) individuais não podem ser utilizados na estimação da demanda. No seu lugar, utilizamos as *dummies* de instituições financeiras (*idif* 1 – *idif* 34) com o intuito de captar as diferenças relativas a custos existentes entre as instituições financeiras. Assumimos que os choques de demanda, tais como preferências do consumidor por determinados bancos por gosto ou costume são captados pelas variáveis de controle (tempo de relacionamento e número de IFs com que o tomador se relaciona).

1.4.3 Especificação econométrica

Para tratar dos problemas de seleção e endogeneidade na estimação da demanda por empréstimos, utilizamos a estimação proposta por Das, Newey e Vella (2003) - DNV. A caracterização semi-paramétrica do modelo que queremos estimar é a **Equação 1.2**, juntamente com as **Equação 1.3-Equação 1.5**:

Equação 1.3: $p(d = 1 / Z) = \Phi(\theta_d Z)$

Equação 1.4: $r = \gamma_j x + \beta_j W + u_r$

Equação 1.5: $T = \gamma_p x + \beta_p W + u_T$

A **Equação 1.3** expressa a seleção amostral e a taxa de juros e o prazo (endógenos) são descritos pela **Equação 1.4** e pela **Equação 1.5**, respectivamente.

A idéia básica para a identificação e estimação deste modelo é a inclusão do *propensity score*, juntamente com os erros estimados u_r e u_T na **Equação 1.2**. $P(d = 1 / Z)$ é desconhecido e por isso deve ser estimado. Neste caso, modelamos a seleção como um *probit* (**Equação 1.3**), onde $\Phi(\cdot)$ é a densidade da normal padrão. Definimos d uma variável binária igual a 1 se o indivíduo comprou e financiou e igual a zero se comprou e pagou à vista. As variáveis explicativas (Z) são *dummies* iguais a 1 para a região de residência do comprador e zero caso contrário.

Na **Equação 1.4** e na **Equação 1.5**, o vetor x é o mesmo da equação do empréstimo. W é o vetor de variáveis que afetam a taxa de juros e o prazo e que podem ser excluídas da equação de empréstimo. A estimação de cada uma das equações é feita por OLS, computando-se os resíduos.

Para a estimação da **Equação 1.2**, a hipótese de Das, Newey e Vella (2003) é que a média condicional do erro, ε , dada a seleção, as variáveis endógenas (*jur, prazo*) e as covariadas x, Z é função somente do *propensity score* e dos erros não observados das variáveis endógenas, ou seja, $E(\varepsilon / x, Z, jur, prazo, d = 1) = \lambda(u_r, u_T, p(d = 1 / Z))$. Podemos reescrevê-la então como:

Equação 1.6: $\ln(emprest) = \gamma x + f(jur, prazo) + \lambda(u_r, u_T, p(d = 1 / Z))$

A **Equação 1.6** é parcialmente linear, onde $f(\cdot)$ é um componente não especificado que queremos estimar (parte não-paramétrica). A estimação do modelo é feita da seguinte forma:

1. Estima-se o *propensity score* $p(d=1/Z)$ para cada uma das 54 regiões (*reg*) na amostra da POF 2002-2003 usando os coeficientes estimados da regressão *probit* (**Equação 1.3**) e imputa-se na base do SCR;
2. Estimam-se as equações de taxa de juros e prazo (**Equação 1.4** e **Equação 1.5**) utilizando as variáveis exógenas e os instrumentos e encontram-se os resíduos (r_{jur}, r_{prazo});
3. Estima-se a equação do empréstimo (**Equação 1.2**), corrigindo para os potenciais vieses de seleção e endogeneidade, inserindo-se o *propensity score* e os resíduos das equações de taxas de juros e prazo (de forma não-paramétrica). A especificação escolhida é a que minimiza o critério do *cross-validation* (CV). O CV é a soma dos quadrados dos erros previstos, onde todas as demais observações são utilizadas para prever cada uma das observações¹¹;
4. Estima-se os desvios-padrão dos coeficientes por *bootstrap*, para considerar o fato de que o *propensity score* é estimado com os dados da POF;
5. Verifica-se qual a resposta do consumidor com relação a variação na taxa de juros e no prazo de pagamento;
6. Divide-se a amostra em duas partes, uma contendo os consumidores mais prováveis de sofrerem a restrição e outra com as menos prováveis para encontrar evidências da presença de restrição ao crédito neste mercado. Refaz-se a estimação com cada um dos subconjuntos da amostra.

¹¹ Por motivo de agilidade no processamento, retiramos duzentas observações por vez, ao invés de apenas uma.

1.5 BASE DE DADOS

Para estimar o modelo utilizamos as bases de dados do SCR – Sistema de Informações de Crédito e do ESTBAN do Banco Central do Brasil¹², da Pesquisa de Orçamentos Familiares – POF 2002-2003 e da base Estatcart de Informações Municipais do IBGE.

A base do SCR é a mais rica fonte de dados que tivemos acesso, pois traz todas as operações iniciadas em maio/2004 na modalidade de Aquisição de Veículos Automotores¹³ - Pessoa Física para os valores de empréstimo superiores a R\$ 5.000,00. A alimentação desta base é feita diretamente pela instituição financeira que concedeu o empréstimo. Apesar de não estar livre de erros de imputação dos dados, a base está isenta de erros de superestimação/subestimação da resposta do entrevistado, que pode ocorrer em dados oriundos de respostas de questionários.

¹² Uma descrição detalhada da base do SCR referente às operações de crédito pré-fixadas iniciadas em mai/2004 é feita por Koyama (2007). Conforme descreve o autor, a modalidade Aquisição de Veículos Automotores – PF apresentou 165.708 operações (12,6% do total), totalizando um volume de empréstimos de R\$ 2.106 milhões (6,13% do total). Para este trabalho excluímos as operações com classificação de risco HH, taxa de juros inferior a 1% a.a., IFs com menos de 20 operações e as com operações concentradas em apenas uma região.

¹³ A base do SCR não inclui as operações de *leasing*. No período utilizado (mai/2004), essas operações não compunham parte significativa da forma de pagamento (vide **Seção 1.2 - Figura 1.3**).

Conforme Almeida (2007) as principais diferenças entre o CDC – Crédito Direto ao Consumidor e o *Leasing* são:

CDC: 1) Prazo: A partir de seis meses, e pode ser quitado a qualquer momento, com desconto se houver antecipação de algumas parcelas ou do total; 2) Propriedade do veículo: O documento do carro fica em nome do comprador, com alienação ao banco que faz o financiamento. Na hora de transferir, a baixa é feita automaticamente após a quitação; e 3) Benefício fiscal e tributação: Não há benefício fiscal. É cobrado IOF (Imposto sobre Operações Financeiras), o que faz a taxa cobrada ser maior.

Leasing: 1) Prazo: Começa a partir de 24 meses, prazo que é exigido também para liquidação do contrato. Não há desconto para antecipação de parcelas, pois se trata de uma prestação de serviço e não de uma operação financeira; 2) Propriedade do veículo: O documento fica em nome da empresa e não do comprador (arrendatário). O veículo não precisa ser lançado na declaração do Imposto de Renda. A transferência é um pouco mais burocrática, pois exige uma carta de autorização e de opção de compra; e 3) Benefício fiscal e tributação: Para pessoa física não há benefício. No caso de pessoa jurídica, há vantagem fiscal, desde que a empresa esteja enquadrada como lucro real. Nesse caso, as prestações são lançadas como despesa, o que ajuda a reduzir o Imposto de Renda.

O financiamento do veículo esta limitado a 100% do seu valor ($\phi_{\max} = 100\%$). A base do SCR traz o valor de $\phi.G$, mas não os de ϕ e G separadamente¹⁴. Por isso não foi possível inferir qual o valor dado de entrada, por exemplo.

Os dados do SCR podem ser divididos em três blocos: as características do contrato, as características do cliente e as características do banco. Com respeito às características do contrato de financiamento temos: valor do empréstimo (*empr*), taxa de juros (*jur*), prazo da operação (*prazo*) e classificação de risco (*risco*). Quanto às características do cliente, temos: tempo de relacionamento com o banco (*relac*), tempo de relacionamento ao quadrado (*relac2*) e o número de bancos com os quais mantém relacionamento (*totif*). Como características do banco, além de identificá-lo (*idif*), sabemos a região da agência bancária que concedeu o empréstimo (*reg*).

Com o intuito de considerar um possível contrato substituto, com os dados do SCR construímos as variáveis *subj1* e *subpl*. Elas representam, respectivamente, a média da taxa de juros e do prazo de pagamento dos contratos ofertados por outros bancos (vide **Tabela 1.6**, para uma descrição mais detalhada).

Consideramos que a variável *renda* é importante para mensurar o efeito riqueza neste mercado. Para a construção desta, ajustamos uma distribuição *lognormal* (μ, σ) das UCs de todos os indivíduos que declararam ter incorrido em despesas de aquisição de veículos para cada uma das regiões, com os dados obtidos da POF – 2002-2003, que foi imputada na base do SCR. Para maiores detalhes da construção dessa variável, vide **Tabela 1.6** e o **APÊNDICE 1.A**, p.50.

Dessa forma, o vetor x é representado pelas variáveis: *risco*, *relac*, *relac2*, *totif*, *renda*, *subj1* e *subpl*.

¹⁴ Uma *proxy* para o valor do automóvel seria o valor da garantia não fidejussória. No entanto, na base que tivemos acesso essa variável apresenta muitos valores *missing* e iguais a zero, não se mostrando adequada para a estimação.

O sistema ESTBAN traz as principais contas do balancete dos bancos para cada agência em cada município brasileiro. Dessa base, computamos os valores do índice *Herfindahl* para o volume de empréstimo para cada município para dez/2003 (*hempr*). Da base Estatcart de Informações Municipais do IBGE obtivemos a relação crédito/PIB (*pot*) e o número de agências bancárias por mil habitantes (*nagencph*). Essas três variáveis foram utilizadas como instrumentos na equação de demanda.

1.5.1 Análise preliminar dos dados

Esta subseção tem como objetivo apresentar de forma descritiva e resumida as principais características da base de dados do SCR utilizada neste trabalho.

Uma operação média de Financiamento de Veículo - PF em mai/2004 possui valor do empréstimo de R\$ 12.890,95, taxa de juros efetiva de 32,99% a.a. e prazo de 32 meses (**Tabela 1.4**). O tempo de relacionamento médio do tomador com a instituição é de 2 anos, sendo que mais de 50% tornou-se cliente no mês da aquisição do veículo. Mais de 90% das operações são consideradas pouco arriscadas, com classificação de risco igual a 2 (igual a “A”, na escala de letras), reflexo do fato de o próprio veículo financiado servir de garantia. A variável renda construída com base na POF 2002-2003 apresenta média de R\$ 3.936,75. O comparativo da distribuição *lognormal* da renda por região dos dados da POF com os valores imputados no SCR é feito no **APÊNDICE 1.A**, p.50.

Tabela 1.4 – Estatística descritiva das variáveis

Variável	Média	Desv. Padr.	1º quartil	Mediana	3º quartil
<i>emprest</i> (R\$)	12.890,95	16.755,27	6.807,15	9.392,51	14.459,79
<i>jur</i> (% a.a.)	32,99	11,81	26,08	32,30	39,04
<i>prazo</i> (em meses)	31,78	9,31	24,33	36,50	36,50
<i>relac</i> (em anos)	2,02	4,78	0	0	1,65
<i>risco</i>	2,07	0,43	2	2	2
<i>renda</i> (R\$)	3.936,75	4.138,53	1.531,91	2.750,55	4.883,12
<i>totif</i>	1,50	0,94	1	1	2
<i>subj1</i>	32,86	3,90	29,58	32,70	34,83
<i>subp1</i>	31,65	1,46	30,73	31,70	32,69

Nº Observações: 143.636. Fontes: **Bacen**, exceto *renda*, *subj1* e *subp1*.

Na **Tabela 1.5**, apresentamos a estatística descritiva das variáveis por faixas de valores de empréstimo. Alguns padrões são evidentes. Quanto menor a taxa de juros e maior o prazo de pagamento, maior é o valor do empréstimo. Quanto maior o tempo de relacionamento do cliente com a instituição financeira, maior o valor do empréstimo. Com relação ao risco, as operações tendem a tornar-se em média um pouco mais arriscadas quanto maior o valor do empréstimo. A variável renda mantém o critério utilizado para a sua imputação: quanto maior o valor do empréstimo, maior o valor da renda. Quanto maior o número de **IFs** com as quais o cliente se relaciona, maior tende a ser o valor tomado.

Tabela 1.5 – Estatística descritiva das variáveis por quartis de valores de empréstimo

Quartil:	1º		2º		3º		4º	
Valor do empréstimo:	até R\$ 6.807,15		de R\$ 6.807,15 a R\$ 9.392,51		de R\$ 9.392,51 a R\$ 14.459,79		a partir de R\$ 14.459,79	
Variável	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
<i>jur</i> (% a.a.)	38,54	10,03	34,40	11,95	30,67	11,36	28,35	11,20
<i>prazo</i> (em meses)	30,47	8,26	32,00	8,27	31,79	9,92	32,85	10,43
<i>relac</i> (em anos)	1,61	4,16	1,80	4,43	2,04	4,88	2,64	5,49
<i>risco</i>	2,06	0,39	2,05	0,40	2,06	0,40	2,11	0,50
<i>renda</i> (R\$)	2.477,49	3.335,11	3.080,35	3.025,21	3.967,12	3.112,98	6.222,03	5.518,96
<i>totif</i>	1,27	0,63	1,34	0,73	1,49	0,87	1,89	1,28

Fonte: SCR – Bacen, exceto *renda*.

A **Tabela 1.6** traz a descrição das variáveis utilizadas no modelo:

Tabela 1.6 – Descrição das variáveis

Variável	Descrição
<i>lemprest</i>	$\ln$ (logaritmo natural) do valor financiado (R\$ mil).
<i>jur</i>	Taxa de juros (% a.a.) do financiamento.
<i>prazo</i>	Prazo do financiamento em meses.
<i>relac</i>	Tempo de relacionamento do tomador do empréstimo com a instituição financeira em anos.
<i>relac2</i>	Variável <i>relac</i> ao quadrado.
<i>risco</i>	Classificação de risco da operação pelo banco, divididas em nove categorias, que varia de “AA” a “H”, sendo a classificação “AA” considerada a menos arriscada e “H” a mais arriscada. Na base, atribuiu-se os seguintes valores: AA=1, A=2, B=3, C=4, D=5, E=6, F=7, G=8 e H=9 ¹⁵ .
<i>renda</i>	<p>$\ln$ da renda. Imputamos a renda da POF 2002-2003 na base do SCR utilizando o seguinte critério: para os indivíduos que na POF declararam despesas com aquisição de automóveis (passeio e utilitário, nacional e importado) e motos, à vista ou prazo (para a própria Unidade de Consumo ou para outra UC), selecionamos a respectiva renda da UC. Para cada uma das 54 regiões, ajustamos uma distribuição <i>lognormal</i> (μ, σ) da renda e a dividimos em quintis, $i = 1, \dots, 5$.</p> <p>Paralelamente, na base do SCR, para cada região, ordenamos os valores do empréstimo de forma crescente e o dividimos em cinco partes iguais, $i = 1, \dots, 5$. A quinta parte q_i, recebe de forma aleatória uma renda com uma distribuição <i>lognormal</i> (μ, σ) sendo que 60% refere-se ao quintil correspondente e 40% ao restante da distribuição. Para evitar que a mesma observação fosse selecionada diversas vezes, geramos a <i>lognormal</i> com cem vezes mais observações que as necessárias para cada região.</p>
<i>dreg1 a dreg54</i>	Variáveis <i>dummies</i> =1 se a agência bancária que concedeu o empréstimo está localizada na região correspondente e =0 caso contrário. Cada uma das 27 UFs foi dividida em duas partes: capital e resto da UF, totalizando 54 regiões.
<i>d</i>	Variável <i>dummy</i> =1 se o indivíduo comprou e financiou o veículo e =0 se comprou e pagou à vista. Esse dado foi extraído da POF 2002-2003 e refere-se aos indivíduos que declararam despesas com aquisição de automóveis (passeio e utilitário, nacional e importado) e motos, à vista e a prazo (para a própria Unidade de Consumo ou para outra UC).
<i>totif</i>	É o número de instituições financeiras que o tomador do empréstimo mantém relacionamento na data-base.
<i>didif1 a didif34</i>	Variáveis <i>dummies</i> =1 para a instituição financeira que concedeu o empréstimo (34 IFs).
<i>subj1</i>	É a <i>proxy</i> da taxa de juros substituta. As variáveis <i>emprest</i> e <i>renda</i> foram divididas em três partes: valores menores ou iguais ao percentil 25, valores maiores ou iguais ao percentil 75 e valores entre esses percentis. Seja a IF j , $j = 1, \dots, 34$; a região i , $i = 1, \dots, 54$; o <i>emprest</i> k , $k = 1, \dots, 3$ e <i>renda</i> l , $l = 1, \dots, 3$. A variável <i>subj1</i> para a observação do tipo (j, i, k, l) é igual a

¹⁵ Não excluímos as classificações de risco acima de 5, pois correspondem a somente 0,5% da amostra.

Variável	Descrição
	média das taxas de juros praticadas pelas IFs $\neq j$, no estrato geográfico i , com <i>emprest</i> k , <i>renda</i> l e <i>relac</i> inferior a 180 dias.
<i>subpl</i>	Análogo à variável <i>subj1</i> , para o prazo.
<i>hemprest</i>	Índice de <i>Herfindahl</i> para o volume de empréstimo calculado para cada município com dados do ESTBAN de dez/2003. O índice de <i>Herfindahl</i> é calculado da seguinte forma: $hempr = \sum_{j=1}^n s_j^2$, onde s_j^2 é o quadrado da proporção do volume de empréstimos do banco j no município.
<i>pot</i>	Relação crédito(2006)/PIB(2004) para cada município, calculados com dados da Base Estatcart de Informações Municipais – 2006, edição 2007.
<i>nagencph</i>	Relação entre o número de agências (2006) e a população (1000 habitantes, 2006) para cada município, calculados com dados da Base Estatcart de Informações Municipais – 2006, edição 2007.
<i>ps</i> , <i>ps2</i> , <i>ps3</i>	<i>Propensity score</i> , <i>propensity score</i> ao quadrado e <i>propensity score</i> ao cubo, respectivamente.
<i>rjur</i> , <i>rjur2</i>	Resíduo da equação da taxa de juros no nível e ao quadrado, respectivamente.
<i>rprazo</i> , <i>rprazo2</i> , <i>rprazo3</i>	Resíduo da equação do prazo de pagamento no nível, ao quadrado e ao cubo, respectivamente.

1.6 RESULTADOS

1.6.1 Amostra completa

Esta subseção apresenta os resultados empíricos considerando a amostra completa, com 143.636 observações. Com o intuito de escolher a especificação para a equação de empréstimo utilizamos o critério que minimiza o *cross-validation* – CV, que é a soma dos quadrados dos erros previstos (no qual se retira 200 observações de cada vez, roda-se a regressão e computam-se os erros previstos das observações excluídas). A **Tabela 1.7** traz o valor do CV para 16 diferentes especificações. Inicialmente, verificamos que o **Modelo 1** sem a inclusão do *propensity score* e dos resíduos das equações de taxas de juros e prazo é claramente rejeitado por esse critério, pois a inclusão dos mesmos diminui o valor do CV (comparando com os **Modelos 2 a 4**).

O *propensity score* apresenta pouca variabilidade (48 para uma amostra com cerca de 140 mil observações) e por isso a sua contribuição para corrigir o problema de seleção da amostra pode não ser significativa. A base de dados utilizada para tratar da seleção é a POF - 2002-2003 e as *dummies* de região foram as únicas variáveis que nos possibilitaram realizar a ligação com a base do SCR. O *propensity score* parece contribuir para minimizar o CV, pois as especificações que não o incluem apresentam os maiores valores do CV (vide **Modelos 1, 3, 5 e 11**). Por isso, de acordo com esse critério, é justificável a inclusão do *propensity score* na especificação. Um segundo motivo que sustenta a sua inclusão é que, nas estimativas, o seu coeficiente é significativo (vide **Tabela 1.8**).

Dentre todas as especificações apresentadas, o **Modelo 9** é o que minimiza esse critério e, portanto, foi o escolhido. Neste modelo, o vetor de variáveis exógenas x é composto pelas variáveis *relac*, *relac2*, *risco*, *lrenda*, *totif*, *subj1* e *subp1*, além das variáveis endógenas *jur* e *prazo*. A não-linearidade dos efeitos da taxa de juros e da maturidade é capturada pela interação entre *jur* e *prazo*, além de *jur2*, *prazo2* e *prazo3*. A especificação torna-se completa com $\lambda(.)$ representado por *rjur*, *rjur2*, *rprazo*, *rprazo2*, *ps*, *ps2* e *ps3*.

Tabela 1.7 – Resultados do critério de *cross-validation*

Modelo	Especificação							CV
	J	P	RJ	RP	PS	I1	I2	
1	1	1	0	0	0	0	0	16.345
2	1	1	0	0	1	0	0	15.711
3	1	1	1	1	0	0	0	15.952
4	2	3	1	1	1	0	0	14.667
5	1	1	2	1	0	3	0	17.582
6	1	1	2	1	3	3	0	16.774
7	2	3	2	2	3	0	0	14.537
8	1	1	1	2	2	2	0	14.888
9	2	3	2	2	3	1	0	14.387
10	3	3	0	2	3	3	1	14.835
11	2	3	2	1	0	2	1	17.006
12	1	1	0	0	3	3	0	15.865
13	2	3	2	2	0	1	0	15.124
14	2	3	2	2	0	0	0	15.284
15	3	3	2	2	3	3	0	15.525
16	3	3	2	2	3	3	1	15.525

Nota: Os valores de juros (**J**), prazo (**P**), resíduo dos juros (**RJ**), resíduo do prazo (**RP**), *propensity score* (**PS**) que são discriminados denominam o grau da variável na especificação da equação do *log* do empréstimo (**Equação 1**). Por exemplo, **J**=3 significa que a especificação contém a variável juros no nível, ao quadrado e ao cubo. **I1**=1 significa que o modelo contém interação entre os termos lineares **J** e **P**; **I1**=2 refere-se as interações entre o termo linear e quadrático de **J** e **P**; **I1**=3 corresponde a **I1**=1 e **I1**=2 mais a interação entre os termos quadráticos de **J** e **P**; **I2**=1 refere-se a inclusão das interações entre os termos lineares **RJ**, **RP** e **PS** (dois a dois).

As duas primeiras colunas da **Tabela 1.8** mostram as estimativas da equação de demanda por empréstimos por OLS (sem considerar endogeneidade nem seleção) e 2SLS (considerando apenas endogeneidade). Todas as variáveis exceto a taxa de juros e o prazo apresentam coeficientes bastante similares. Tanto o coeficiente do juros quanto do prazo subestimam os valores para mais próximos a zero na estimação por OLS.

Tabela 1.8 – Resultados da estimativa da Equação 1

Variável	OLS ^(a)		2SLS ^{(a), (b)}		DNV - Modelo 9 ^(c)	
	Coeficiente	Desvio-padrão	Coeficiente	Desvio-padrão	Coeficiente	Desvio-padrão
<i>jur</i>	-0,004883	0,000656	-0,013268	0,000333	-0,012090	0,000687
<i>prazo</i>	0,001430	0,000118	0,015586	0,000988	0,037610	0,002118
<i>relac</i>	0,006759	0,000300	0,007807	0,000329	0,007110	0,000328
<i>relac2</i>	-0,005553	0,000517	-0,006334	0,000542	-0,006181	0,000573
<i>risco</i>	0,063116	0,003161	0,061407	0,003029	0,056081	0,003031
<i>lrenda</i>	0,056721	0,001382	0,052370	0,001407	0,052614	0,001202
<i>totif</i>	0,072771	0,001513	0,070451	0,001433	0,071811	0,001271
<i>subj1</i>	-0,095351	0,000471	-0,086196	0,000609	-0,082616	0,000393
<i>subp1</i>	0,045042	0,001588	0,041699	0,001299	0,031952	0,000930
<i>iter11=jur*prazo</i>	-	-	-	-	-0,000308	0,000009
<i>jur2</i>	-	-	-	-	0,000125	0,000008
<i>prazo2</i>	-	-	-	-	-0,001192	0,000059
<i>prazo3</i>	-	-	-	-	0,000018	0,000001
<i>rjur</i>	-	-	-	-	0,007338	0,000285
<i>rprazo</i>	-	-	-	-	-0,008580	0,000846
<i>rjur2</i>	-	-	-	-	-0,000122	0,000009
<i>rprazo2</i>	-	-	-	-	-0,000272	0,000030
<i>ps</i>	-	-	-	-	11,473810	0,577079
<i>ps2</i>	-	-	-	-	-20,006670	1,016304
<i>ps3</i>	-	-	-	-	10,433660	0,589919
<i>constante</i>	10,042360	0,068917	9,727548	0,068405	8,338866	0,111531
<i>Dummies de região</i>	sim		sim		não	

A variável dependente é *lemprest*. N° de observações: 143.636. Coeficientes significativos a 1%.

^(a) Desvio-padrão robusto à heterocedasticidade.

^(b) Instrumentos: *nagenph*, *hempr*, *pot*.

^(c) Desvio-padrão calculado por *bootstrap*, com 200 replicações.

A terceira coluna da **Tabela 1.8** apresenta a estimativa da equação de demanda por empréstimos pela forma não-paramétrica proposta por DNV (2003) para a especificação número 9 da **Tabela 1.7**. O **APÊNDICE 1.B**, p.52 deste trabalho traz o *printout* do *Stata* das estimativas das **Equação 1.3-Equação 1.5** do modelo DNV.

Observa-se que a estimativa por 2SLS em relação ao DNV, excetuando-se o coeficiente do prazo, apresenta valores bastante similares. Os coeficientes da taxa de juros e do prazo, da sua iteração, do *propensity score*, dos resíduos da equação da taxa de juros e prazo de pagamento, bem como da inclusão dos termos ao quadrado e ao cubo são todos significativos na

estimação por DNV. Este é um indício de que tanto o problema de seleção quanto o de endogeneidade devem ser tratados para estimar a demanda por financiamento de veículos.

A variável *renda* que foi imputada na base do SCR apresenta coeficiente positivo e significativo: quanto maior a renda, maior o valor do empréstimo. Os coeficientes das variáveis *relac* (positivo) e *relac2* (negativo) indicam que quanto maior o tempo de relacionamento do cliente com a instituição, maior tende a ser o valor do empréstimo, a uma taxa decrescente.

A maioria das operações de crédito é classificada como sendo de baixo risco (classificação igual a “A”) e, na estimação, o seu coeficiente é positivo e significativo. Nos financiamentos, como o próprio veículo serve de garantia, uma classificação mais arriscada do empréstimo pode estar relacionada com o risco do próprio cliente. E um cliente mais arriscado desejaria tomar um valor maior emprestado.

Totif apresenta sinal positivo e significativo indicando que quanto maior o número de instituições financeiras com a qual o tomador se relaciona, maior o volume que pode obter financiado. Este resultado é compatível com a explicação de que relacionar-se com diversos bancos diminui o custo de procura por contratos (*jur,matur*) que melhor atendam às suas necessidades.

As variáveis *subj1* e *subp1* representam um possível contrato de financiamento substituto. Elas representam os valores médios da taxa de juros e do prazo de financiamento dos demais contratos realizados por outras IFs para valores de financiamento e renda similares, numa dada região e tempo de relacionamento inferior a 180 dias. Se os efeitos da assimetria de informação diminuem com o tempo de relacionamento, quando um cliente de longa data de uma IF buscar outro banco, será tratado por este como um cliente novo. A princípio, se esperaria que quanto maior a taxa de juros desse bem (*subj1*) e menor o seu prazo (*subp1*), mais vantajoso seria o verdadeiro contrato e, portanto, maior o valor do empréstimo. No entanto, as estimativas dessas variáveis apresentam sinais opostos ao esperado pela teoria. Por um lado, se na média os demais contratos oferecidos são mais atraentes, mas ainda assim o consumidor prefere o seu contrato, significa que o mesmo pode estar restrito a esse crédito ou

o *switching cost* é muito elevado. Também é possível que essas variáveis não estejam concebidas adequadamente.

A magnitude das elasticidades da taxa de juros e do prazo de pagamento indica a intensidade da resposta da demanda por financiamento de veículos. A **Tabela 1.9** traz os valores das elasticidades da taxa de juros e do prazo de pagamento calculados para o valor médio das variáveis para especificação da terceira coluna da **Tabela 1.8**. Ambos os valores apresentam os sinais preditos pela teoria:

Tabela 1.9 – Elasticidade da demanda por financiamento de veículos – Modelo 9

taxa de juros		prazo	
-0,4497	(0,0099)	+0,3293	(0,0268)

Coeficientes significativos a 1%.

Desvio-padrão calculado por *bootstrap* entre parênteses (200 replicações).

A elasticidade da taxa de juros (0,45) é interpretada da seguinte maneira: se a taxa de juros diminuísse em 1%, a demanda por empréstimo aumentaria em 0,45%. De forma análoga, a elasticidade do prazo (+0,33) sugere que se aumentasse o vencimento da operação em 1%, a demanda por empréstimos aumentaria 0,33%. Esses resultados indicam que o consumidor responde tanto à taxa de juros quanto ao prazo de pagamento.

O fato de a elasticidade da taxa de juros ser superior à do prazo não permite concluir que neste mercado não está presente a restrição ao crédito. Conforme argumentado, se ambas variáveis são importantes na definição do limite de crédito, então o consumidor pode responder mais a uma ou à outra. Pelos dados do mercado de financiamento de veículos da **Seção 1.2**, verificamos que apesar de esta modalidade apresentar as taxas de juros mais baixas, ainda assim é bastante elevada se comparada com a SELIC. Com relação ao prazo de pagamento, a base do SCR apresenta uma média de 32 meses, enquanto o mercado disponibilizava prazos de até 60 meses. Na subseção a seguir, procedemos à divisão da amostra para tentar encontrar maiores indícios a respeito deste problema.

1.6.2 Subdivisão da amostra

Para encontrar evidências sobre a importância da restrição ao crédito no mercado de financiamentos de veículos, repetiremos a mesma especificação da subseção anterior (**Modelo 9**), desta vez com a amostra dividida em duas partes, uma contendo os indivíduos mais prováveis e a outra com os menos prováveis de sofrer a restrição. Conforme já mencionado, esta é uma abordagem frequentemente utilizada pela literatura (Bernanke, 1984; Hayashi, 1985; Zeldes, 1989; AGK, 2008) pois em geral, os indivíduos que se defrontam com a restrição não são diretamente observados.

Para montar os subgrupos, dentre as variáveis disponíveis, selecionamos as mais prováveis de identificar a restrição de crédito. São elas: o valor do empréstimo, da prestação, da renda e comprometimento da renda. O valor da prestação foi calculado multiplicando-se o valor do empréstimo pelo Fator de Recuperação de Capital $\left(= \frac{(1 + \text{juros})^{\text{prazo}} \cdot \text{juros}}{(1 + \text{juros})^{\text{prazo}} - 1} \right)$. A parcela de comprometimento da renda foi calculada dividindo-se o valor da prestação pela renda imputada. Para cada ordenação dessas quatro variáveis, dividimos a amostra em duas partes, uma com os maiores e outra com os menores valores, resultando em oito subgrupos. Para cada uma delas, estimamos utilizando a especificação do **Modelo 9** e calculamos as elasticidades. *A priori* espera-se que os subgrupos que sofrem restrição ao crédito sejam aqueles que contêm os maiores valores de empréstimo, prestação e comprometimento da renda e menores valores de renda.

A **Tabela 1.10** apresenta as estimativas das elasticidades calculadas no valor médio das variáveis¹⁶. Em todos os casos, os sinais das variáveis estão de acordo com o esperado pela teoria. Com maior ou menor intensidade, verifica-se que todos os subgrupos respondem tanto à taxa de juros quanto ao prazo de pagamento, sendo que o subgrupo com os menores valores de empréstimo apresentam os menores valores em módulo. De modo geral, as subdivisões mais prováveis de sofrerem restrição ao crédito respondem mais à variação no prazo de pagamento bem como à taxa de juros com relação à sua contraparte.

¹⁶ Para efeito de comparação, calculamos as elasticidades para todas as observações da amostra (total e dividida) e computamos a média, vide **APÊNDICE 1.C**, p.55. O desvio-padrão apresentado é o da amostra correspondente. Verifica-se que o coeficiente das elasticidades calculada no ponto médio e a média das elasticidades não apresentam diferenças significativas. No entanto, a média das elasticidades tendem a apresentar o desvio-padrão bastante elevado devido a observações que geram valores muito alto ou muito baixo.

Tabela 1.10 – Elasticidades para o Modelo 9 para as diversas subamostras

Divisão por:	Valor da prestação		Renda		Valor do empréstimo		Comprometimento da renda	
Variável	1/3 inferior	1/3 superior	1/3 inferior	1/3 superior	1/2 inferior	1/2 superior	<= 0,25	>0,25
<i>jur</i>	-0,2756 (0,0082)	-0,2884 (0,0129)	-0,4766 (0,0189)	-0,4214 (0,0241)	-0,1130 (0,0055)	-0,3271 (0,0150)	-0,3965 (0,0082)	-0,5006 (0,0240)
<i>prazo</i>	0,9312 (0,0252)	1,0470 (0,0260)	0,4687 (0,0451)	0,1704 (0,0426)	0,0456 (0,0125)	0,2473 (0,0313)	0,3044 (0,0279)	1,0009 (0,0321)
<i>Nº obs:</i>	47.879	47.879	47.879	47.879	71.818	71.818	97.746	45.890

Em azul, o subgrupo mais provável de sofrer restrição. Todos os coeficientes significativos a 1%.

Desvio-padrão entre parênteses e calculado por *bootstrap*, com 200 replicações.

A divisão da amostra pelo valor da prestação parece não influir no valor da elasticidade da taxa de juros, que é de -0,28 para ambos subgrupos. A amostra com os maiores valores de prestação, no entanto, apresenta maior sensibilidade quanto ao prazo de pagamento.

Se separarmos a amostra pela renda, ambas subamostras apresentam novamente elasticidade similar da taxa de juros, sendo -0,48 para os de renda mais baixa contra -0,42 para o de renda mais alta. Se compararmos a resposta quanto à variação no prazo de pagamento, verifica-se que o primeiro grupo apresenta valor 2,7 vezes maior. Ademais, se analisarmos apenas os consumidores com maior renda, constata-se que os mesmos respondem mais a taxa de juros que ao prazo de pagamento.

Dividindo a amostra total pelo valor do empréstimo obtemos os seguintes resultados: os consumidores que tomam os maiores empréstimos respondem mais a variação na taxa de juros e no prazo do que aqueles que tomam menos. Em termos proporcionais, a elasticidade da taxa de juros e do prazo de pagamento é cerca 3 e 5,5 vezes maior, respectivamente.

A parcela de comprometimento da renda apresenta mais evidências. Os consumidores que comprometem menos do orçamento apresentam elasticidade da taxa de juros maior que a do prazo de pagamento (0,40 contra 0,30). Por outro lado, aqueles que comprometem mais de 25% da renda respondem mais ao prazo de pagamento do que à taxa de juros (1,0 contra 0,5). Os que comprometem mais da renda respondem mais à taxa de juros (25% mais) e ao prazo de pagamento (3,3 vezes maior) do que o restante da amostra. Colocando de outra forma, esse

resultado indica que se a taxa de juros diminuir 1%, a demanda dos que comprometem menos da renda aumentaria 0,4% e dos que comprometem mais, 0,5%. Analogamente, se o prazo de pagamento aumentasse 1%, a demanda desses grupos aumentaria 0,3% e 1%, respectivamente.

Para comparar, apresentamos alguns dos resultados encontrados por AGK (2008) para os dados com consumidores norte-americanos (derivada da taxa de juros e da maturidade):

- 1) amostra completa: taxa de juros = -0,09* (0,25)
maturidade = +7,38 (1,33)
- 2) amostra dividida pela renda (para o grupo de renda per capita mais baixa e mais alta, respectivamente): taxa de juros = -0,31 (0,17) e -1,40 (0,82)
maturidade = +6,17 (1,72) e +0,21* (2,10)

Desvio-padrão entre parênteses. * Coeficiente não significativo.

Os autores não rejeitam a hipótese de que a sensibilidade da demanda com respeito à taxa de juros é igual a zero para a amostra completa com os dados americanos. Ademais, os consumidores com menor renda apresentam menos sensibilidade à taxa de juros e mais à maturidade em relação ao grupo de maior renda.

No nosso caso, o fato de os subgrupos mais prováveis de se defrontarem com a restrição ao crédito apresentarem maior elasticidade do prazo, com relação ao subgrupo menos provável, é um indício de que neste mercado está presente a restrição. Ao contrário do encontrado por AGK (2008), obtemos evidência de que os consumidores mais prováveis de sofrer a restrição são também mais sensíveis à taxa de juros. Essa diferença pode ser explicada pelo fato de que no Brasil as taxas são elevadas (se comparadas com os Estados Unidos) e, portanto, uma pequena alteração modificaria sensivelmente os valores das prestações. E por isso, é possível que tanto a taxa de juros e quanto o prazo de pagamento influam no valor financiado e, particularmente, o grupo dos que sofrem restrição apresenta maior sensibilidade a ambas as variáveis.

Conclui-se, portanto que, utilizando-se tanto a amostra completa como as subdivisões entre grupos mais e menos prováveis de sofrer a restrição, os consumidores respondem tanto à taxa de juros quanto ao prazo de pagamento. Na modalidade analisada, apesar de o próprio veículo

financiado servir de garantia da operação, a divisão da amostra nos traz evidências de que neste mercado está presente a restrição de crédito, principalmente entre aqueles que comprometem parcela maior da renda.

1.7 CONCLUSÃO

O objetivo deste capítulo é verificar empiricamente se os consumidores se defrontam com restrição no mercado de financiamento de veículos. A motivação para este estudo advém do fato de que, no Brasil, existe a percepção de que as pessoas se importam menos com as taxas de juros embutidas nas prestações e mais se o valor da prestação é compatível com o orçamento mensal. Neste caso, espera-se que a elasticidade da demanda com relação ao prazo de pagamento seja superior à da taxa de juros. Se isto ocorre, então é possível que os agentes se defrontem com restrição ao crédito.

Entretanto, essa relação entre as elasticidades só é válida se o estabelecimento de um maior limite de crédito está atrelado basicamente a um maior prazo de pagamento. Se a taxa de juros também é capaz de influir nesse valor, então uma maior sensibilidade a esta variável não contradiz a presença da restrição ao crédito.

Para a estimação da equação de demanda por empréstimo tratamos dos problemas de endogeneidade e seleção amostral, utilizando a metodologia proposta por Das, Newey e Vella (2003).

No teste empírico, foram utilizados os dados do Sistema de Informações de Crédito – SCR do Banco Central do Brasil para as operações iniciadas em maio/2004. Escolheu-se o financiamento de veículos para pessoas físicas, pois: 1) esta base inclui dados individualizados e detalhados de todos os contratos cujo valor é superior R\$ 5.000,00 e, por isso, essa modalidade é mais provável de nela estar presente; e 2) por comprometer parcela significativa da renda, é provável que o consumidor responda a variações na taxa de juros e/ou no prazo de pagamento.

A verificação da existência da restrição ao crédito foi efetuada dividindo-se a amostra em duas partes, uma contendo os consumidores mais prováveis e outro com os menos prováveis de ter este problema.

O resultado encontrado é: os consumidores respondem tanto a variações nas taxas de juros quanto a variações no prazo de pagamento, com elasticidades estimadas para a amostra

completa de -0,45 e 0,33, respectivamente. Este padrão se mantém com a amostra dividida. Os subgrupos mais prováveis de ter restrição são mais sensíveis ao prazo de pagamento e à taxa de juros do que o grupo menos provável. A nosso ver, esse resultado não é contraditório com a presença de restrição se a variação do limite de crédito pode advir de mudanças de uma ou ambas as variáveis. Como no Brasil as taxas de juros são bastante elevadas, uma pequena diminuição deve alterar sensivelmente os valores das prestações.

Dessa forma, conclui-se que no mercado de financiamento de automóveis os consumidores defrontam-se com restrição ao crédito e, ao contrário da percepção comum, eles também respondem a variações nas taxas de juros. Um estudo similar para outras modalidades de financiamento é sugestão para pesquisas futuras.

1.8 REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, Edson Pinto de. Movido a Crédito. + **dinheiro Folha de S. Paulo**. 30 de julho de 2007.

ANEF. Disponível em: <www.anef.com.br>. Acesso em 02/10/2007.

ATTANASIO, Orazio P.; GOLDBERG, Pinelopi K.; KYRIAZIDOU, Ekaterine. *Credit constraints in the market for consumer durables: evidence from micro data of car loans*. **International Economic Review**. Vol. 49, N° 2. May, 2008.

BERNANKE, Ben S. *Permanent income, liquidity, and expenditure on automobiles: evidence from panel data*. **The Quarterly Journal of Economics**. Vol. 99, N° 3, pp. 587-614. Aug., 1984.

DAS, Mitali; NEWHEY, Whitney K.; VELLA, Francis. *Nonparametric estimation of sample selection models*. **Review of Economic Studies**. 70, 33-58. 2003.

FENABRAVE. Disponível em: <www.fenabreve.com.br>. Acesso em 28/09/2007.

GROSS, David B.; SOULELES, Nicholas S. *Do liquidity constraints and interest rates matter for consumer behavior? evidence from credit card data*. **NBER Working Paper 8314**. Jun., 2001.

HAYASHI, Fumio. *The effect of liquidity constraints on consumption: a cross-sectional analysis*. **The Quarterly Journal of Economics**. Vol. 100, N° 1, pp. 183-206. Feb., 1985.

JAPPELLI, Tullio. *Who is credit constrained in the U.S. economy?* **The Quarterly Journal of Economics**. Vol. 105, N° 1, pp. 219-234. Feb., 1990.

JAPPELLI, Tullio; PISCHKE, Jorn-Steffen; SOULELES, Nicholas S. *Testing for liquidity constraints in Euler equations with complementary data sources*. **The Review of Economics and Statistics**. Vol. 80, N° 2, pp. 251-262. May, 1998.

KARLAN, Dean S.; ZINMAN, Jonathan. *Credit elasticities in less-developed economies: implications for microfinance*. **American Economic Review**. 98:3, 1040-1068. 2008.

KOYAMA, Sérgio Mikio. **Empréstimos bancários e operações de redesconto: um estudo sobre modelos de demanda para instituições financeiras**. São Paulo, 2007. Tese (Doutorado em Economia) - Programa de Pós-Graduação em Economia, Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo.

NAKANE, Márcio I.; ALENCAR, Leonardo S. e KANCZUK, Fábio. Demanda por Serviços Bancários. In: FIUZA, Eduardo P.S.; MOTTA, Ronaldo Seroa da (Coord. Téc.). **Métodos quantitativos em defesa da concorrência e regulação econômica**. Tomo 1. Rio de Janeiro, IPEA, 2006.

NEVO, Aviv. *Measuring market power in the ready-to-eat cereal industry*. **Econometrica**. Vol. 69, N° 2, pp. 307-342. Mar., 2001.

NEVO, Aviv. *A practioner's guide to estimation of random-coefficients logit models of demand*. **Journal of Economics & Management Strategy**. Vol. 9, N° 4, pp. 513-548. 2000.

ZELDES, Stephen P. *Consumption and liquidity constraints: an empirical investigation*. **Journal of Political Economy**. Vol. 97, N° 2, pp. 305-346. 1989.

APÊNDICE 1.A

Variável Renda

Para construir a variável renda, utilizamos os dados da POF 2002-2003. Para cada região, selecionamos os indivíduos que declararam ter incorrido em *Despesas com Aquisição de Veículos no Período de Referência de 12 meses* (do tipo automóvel de passeio e utilitário, nacional ou importado e motos), a prazo para a própria Unidade de Consumo - UC ou para Outra UC. Para esses indivíduos, selecionamos a respectiva renda da UC, ajustamos uma distribuição *lognormal* (μ, σ) e a dividimos em quintis.

Paralelamente, no SCR, para cada região, o valor do empréstimo foi ordenado de forma crescente e também foi dividido em quintis. Cada quinta parte recebeu de forma aleatória uma renda com distribuição *lognormal* (μ, σ) sendo que 60% refere-se ao quintil correspondente (da distribuição ajustada da POF) e 40% aos demais quintis. Para diminuir a probabilidade de que a mesma observação fosse selecionada diversas vezes, a *lognormal* foi gerada com cem vezes mais observações do que o número necessário para cada região.

A título de comparação, a **Tabela A - 1.1** traz os parâmetros das distribuições *lognormal* da renda da POF 2002-2003 e das imputadas na base do SCR, por região.

Tabela A - 1.1 – Parâmetros da distribuição *lognormal* da Renda

Região		POF 2002-2003			Imputado no SCR		
		n° obs.	média (μ)	var (σ)	n° obs.	média (μ)	var (σ)
1	Porto Velho	10.202	7,92645	0,82014	236	7,97311	0,79875
2	RO, exceto Porto Velho	28.571	7,71612	0,79093	371	7,72575	0,79061
3	Rio Branco	3.955	8,02439	0,77446	199	8,04087	0,73792
4	AC, exceto Rio Branco	1.908	7,53149	0,64061	21	7,58884	0,50487
5	Manaus	15.828	7,80328	0,94635	1082	7,78516	0,93299
6	AM exceto Manaus	8.819	7,61373	0,78367	95	7,66748	0,79017
7	Boa Vista	1.990	8,00475	0,82197	133	8,05914	0,81538
8	RR, exceto Boa Vista	1.267	7,37319	0,45389	16	7,45359	0,42071
9	Belém	14.029	8,33728	0,61312	1.047	8,32958	0,60531
10	PA, exceto Belém	53.760	7,53537	0,71185	788	7,54355	0,70636
11	Macapá	4.075	8,09407	0,92659	241	8,15120	0,90036
12	AM, exceto Macapá	1.435	7,67594	0,49610	44	7,66201	0,45506
13	Palmas	2.074	7,68784	0,98824	174	7,76652	0,97828
14	TO, exceto Palmas	17.536	7,57936	1,08063	381	7,58628	1,07495
15	São Luís	6.806	8,19219	0,52359	864	8,18545	0,51569
16	MA, exceto São Luís	44.671	7,55700	0,80189	580	7,56111	0,80366

Região		POF 2002-2003			Imputado no SCR		
		n° obs.	média (μ)	var (σ)	n° obs.	média (μ)	var (σ)
17	Teresina	21.730	8,14314	0,80986	545	8,15153	0,82139
18	PI, exceto Teresina	30.399	7,26695	0,80053	241	7,31630	0,77787
19	Fortaleza	50.395	8,20841	0,72123	1.928	8,19809	0,72506
20	CE, exceto Fortaleza	60.954	7,18585	1,01208	660	7,21214	1,03702
21	Natal	14.448	8,23624	1,11699	848	8,22956	1,11109
22	RN, exceto Natal	26.104	7,55428	0,95004	577	7,55732	0,95356
23	João Pessoa	15.800	7,80162	0,88437	733	7,80078	0,88047
24	PB, exceto João Pessoa	28.396	7,22957	1,17232	739	7,22028	1,16634
25	Recife	37.611	7,85422	0,75487	1.672	7,84004	0,75734
26	PE, exceto Recife	77.008	7,29497	0,75755	2.123	7,28441	0,75691
27	Maceió	12.789	8,06049	1,20773	675	8,06197	1,21767
28	AL, exceto Maceió	10.588	7,23630	1,20916	347	7,24225	1,23595
29	Aracaju	10.771	8,24106	0,83316	482	8,26519	0,85555
30	SE, exceto Aracaju	11.351	6,99735	1,24397	206	7,05773	1,21325
31	Salvador	49.274	8,59324	0,66320	1.990	8,59018	0,67012
32	BA, exceto Salvador	117.190	7,37709	1,08911	2.094	7,35728	1,08409
33	Belo Horizonte	70.788	8,12708	0,87230	2.835	8,13641	0,87186
34	MG, exceto Belo Horizonte	305.018	7,69055	0,84865	9.555	7,68575	0,84560
35	Vitória	19.683	8,73408	0,69508	422	8,74703	0,69001
36	ES, exceto Vitória	79.310	7,56971	0,95437	1.519	7,55334	0,93411
37	Rio de Janeiro	136.112	8,36075	0,67133	5.670	8,35744	0,66636
38	RJ, exceto Rio de Janeiro	117.878	7,64992	0,93211	4.875	7,64598	0,91676
39	São Paulo	357.829	8,24303	0,70709	16.052	8,24439	0,70814
40	SP, exceto São Paulo	1.007.338	7,95533	0,79552	37.282	7,95679	0,79737
41	Curitiba	81.724	7,96106	0,75826	3821	7,96681	0,75469
42	PR, exceto Curitiba	235.411	7,66819	0,79717	8594	7,66521	0,79521
43	Florianópolis	26.427	7,70912	0,64323	971	7,70298	0,63086
44	SC, exceto Florianópolis	204.101	7,60528	0,78641	8.896	7,60684	0,78926
45	Porto Alegre	50.940	8,02601	0,88026	1.784	8,01161	0,88294
46	RS, exceto Porto Alegre	320.125	7,67001	0,78243	7.811	7,66315	0,77508
47	Campo Grande	31.829	7,93628	1,04462	1.180	7,93533	1,02790
48	MS, exceto Campo Grande	37.290	7,46322	0,79258	1.427	7,45472	0,79019
49	Cuiabá	10.602	7,69165	1,15898	807	7,70464	1,17505
50	MT, exceto Cuiabá	55.122	7,37843	0,82062	1.792	7,37078	0,82032
51	Goiânia	44.471	7,90740	0,88939	1.806	7,89505	0,88493
52	GO, exceto Goiânia	80.506	7,54026	0,98666	2.751	7,56009	1,00168
53	Brasília	8.143	8,98130	0,66842	1.700	8,96662	0,66713
54	DF, exceto Brasília	67.332	8,30888	0,91593	2.764	8,32557	0,92709
Brasil		4.139.713	7,83670	0,88215	146.446	7,89927	0,87247

* O número de observações diminui para 143.636 devido a valores *missing* das variáveis *hempr*, *nagencph*, *size* e *subj1*.

APÊNDICE 1.B

Printout do Stata

Estimação da Equação 1.3

```
. xi: probit d i.reg [fweight=R0000380]
i.reg          _Ireg_1-54          (naturally coded; _Ireg_1 omitted)

Iteration 0:    log likelihood = -5251542.5
Iteration 1:    log likelihood = -5153341.8
Iteration 2:    log likelihood = -5153172.9
Iteration 3:    log likelihood = -5153172.9

Probit regression                                Number of obs   =    7617869
                                                LR chi2(53)     =   196739.20
                                                Prob > chi2     =    0.0000
Log likelihood = -5153172.9                    Pseudo R2      =    0.0187
```

	d	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
	_Ireg_2	-.6298835	.011963	-52.65	0.000	-.6533305	-.6064364
	_Ireg_3	-.1742015	.0194973	-8.93	0.000	-.2124155	-.1359875
	_Ireg_4	-.1577411	.0258639	-6.10	0.000	-.2084334	-.1070488
	_Ireg_5	-.1518141	.0135938	-11.17	0.000	-.1784574	-.1251708
	_Ireg_6	.2408811	.0169777	14.19	0.000	.2076055	.2741567
	_Ireg_7	-1.26605	.0183771	-68.89	0.000	-1.302069	-1.230032
	_Ireg_8	-1.04512	.022958	-45.52	0.000	-1.090117	-1.000123
	_Ireg_9	-.3718623	.0134459	-27.66	0.000	-.3982158	-.3455088
	_Ireg_10	-.5797922	.0114851	-50.48	0.000	-.6023025	-.5572819
	_Ireg_11	-.1454983	.0194596	-7.48	0.000	-.1836385	-.107358
	_Ireg_12	-.0400851	.0304688	-1.32	0.188	-.099803	.0196327
	_Ireg_13	-.6379341	.0214745	-29.71	0.000	-.6800234	-.5958448
	_Ireg_14	-.4186736	.0128987	-32.46	0.000	-.4439546	-.3933927
	_Ireg_15	-.4290655	.0155412	-27.61	0.000	-.4595257	-.3986054
	_Ireg_16	-.3530679	.0117498	-30.05	0.000	-.376097	-.3300388
	_Ireg_17	.3626564	.0139934	25.92	0.000	.3352299	.3900829
	_Ireg_18	-.4646061	.0120381	-38.59	0.000	-.4882003	-.441012
	_Ireg_19	-.0961852	.0118334	-8.13	0.000	-.1193782	-.0729922
	_Ireg_20	-.5878166	.0114103	-51.52	0.000	-.6101802	-.5654529
	_Ireg_21	-.0813976	.0139893	-5.82	0.000	-.108816	-.0539791
	_Ireg_22	-.4168134	.0122704	-33.97	0.000	-.4408629	-.3927639
	_Ireg_23	-.0094174	.0139155	-0.68	0.499	-.0366913	.0178566
	_Ireg_24	-.6315776	.011968	-52.77	0.000	-.6550345	-.6081206
	_Ireg_25	-.1115721	.0121291	-9.20	0.000	-.1353446	-.0877995
	_Ireg_26	-.1558878	.0114769	-13.58	0.000	-.178382	-.1333935
	_Ireg_27	-.3264205	.013765	-23.71	0.000	-.3533995	-.2994415
	_Ireg_28	-.2518367	.0144843	-17.39	0.000	-.2802254	-.223448
	_Ireg_29	-.1441892	.0147228	-9.79	0.000	-.1730454	-.1153329
	_Ireg_30	-.7503849	.0132455	-56.65	0.000	-.7763456	-.7244243
	_Ireg_31	-.7645787	.0114509	-66.77	0.000	-.7870221	-.7421353
	_Ireg_32	-.6362433	.0111413	-57.11	0.000	-.6580799	-.6144068
	_Ireg_33	-.3262708	.0114435	-28.51	0.000	-.3486997	-.3038419
	_Ireg_34	-.6345825	.0109721	-57.84	0.000	-.6560874	-.6130775
	_Ireg_35	.0538583	.0134992	3.99	0.000	.0274003	.0803163
	_Ireg_36	-.3216704	.0113847	-28.25	0.000	-.343984	-.2993569

```

_Ireg_37 | -.5136594 .0111273 -46.16 0.000 -.5354686 -.4918503
_Ireg_38 | -.5607119 .0111564 -50.26 0.000 -.582578 -.5388458
_Ireg_39 | -.4171384 .0109736 -38.01 0.000 -.4386462 -.3956306
_Ireg_40 | -.3450335 .010906 -31.64 0.000 -.3664089 -.3236582
_Ireg_41 | -.0569299 .0114921 -4.95 0.000 -.0794541 -.0344057
_Ireg_42 | -.459221 .0110242 -41.66 0.000 -.480828 -.437614
_Ireg_43 | .380528 .0135332 28.12 0.000 .3540035 .4070526
_Ireg_44 | -.1260517 .0111058 -11.35 0.000 -.1478186 -.1042848
_Ireg_45 | .0347352 .0119343 2.91 0.004 .0113444 .0581261
_Ireg_46 | -.2978777 .0109986 -27.08 0.000 -.3194345 -.2763208
_Ireg_47 | .0784203 .0125945 6.23 0.000 .0537354 .1031051
_Ireg_48 | -.5910027 .011741 -50.34 0.000 -.6140147 -.5679908
_Ireg_49 | -.3249771 .0142938 -22.74 0.000 -.3529924 -.2969618
_Ireg_50 | -.3124924 .0116108 -26.91 0.000 -.3352491 -.2897356
_Ireg_51 | -.2622178 .0118197 -22.18 0.000 -.2853839 -.2390517
_Ireg_52 | -.8822775 .011196 -78.80 0.000 -.9042212 -.8603338
_Ireg_53 | -.2993539 .0152561 -19.62 0.000 -.3292554 -.2694525
_Ireg_54 | -.1196734 .0115833 -10.33 0.000 -.1423762 -.0969706
_cons | .5134799 .010865 47.26 0.000 .4921849 .5347749

```

```

. predict ps
(option p assumed; Pr(d))
. tabstat ps, statistics(mean) by(reg)
Summary for variables: ps

```

```

by categories of: reg

```

reg	mean	reg	mean	reg	mean
1	.6961921	19	.6617686	37	.4999284
2	.4536664	20	.4703712	38	.4811642
3	.6328	21	.6671592	39	.5383753
4	.6389819	22	.5385044	40	.5668839
5	.6411991	23	.6928913	41	.6760027
6	.7746838	24	.4529951	42	.5216356
7	.225854	25	.6561241	43	.8143412
8	.2974877	26	.6396757	44	.6507804
9	.556309	27	.5741929	45	.7082279
10	.4735646	28	.6032017	46	.5853511
11	.6435565	29	.6440445	47	.7230413
12	.6820342	30	.4063652	48	.4691038
13	.4504778	31	.4008689	49	.5747588
14	.5377656	32	.4511472	50	.5796458
15	.5336365	33	.5742517	51	.5991943
16	.5637217	34	.4518049	52	.3561393
17	.809522	35	.7147578	53	.5847756
18	.5194901	36	.5760543	54	.653138
-----+-----					
				Total	.5479152

Estimação da Equação 1.4 (*dummies* de região e IF não reportados)

i.reg	_Ireg_1-54	(naturally coded; _Ireg_1 omitted)				
i.idif	_Iidif_1-34	(naturally coded; _Iidif_1 omitted)				
Source	SS	df	MS	Number of obs = 143636		
Model	6247591.78	96	65079.0811	F(96,143539) = 678.31		
Residual	13771614.2143539	95	9433619	Prob > F = 0.0000		
Total	20019206143635	139	375542	R-squared = 0.3121		
				Adj R-squared = 0.3116		
				Root MSE = 9.7951		
jur	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
subj1	.5456844	.0112162	48.65	0.000	.5237009	.5676679
subp1	-.415344	.0314826	-13.19	0.000	-.4770493	-.3536387
relac	-.1469734	.0084654	-17.36	0.000	-.1635654	-.1303814
relac2	.1120795	.0160089	7.00	0.000	.0807024	.1434566
risco	1.36624	.0726928	18.79	0.000	1.223764	1.508716
lrenda	-.1871445	.0355897	-5.26	0.000	-.2568996	-.1173894
totif	.0934756	.0286928	3.26	0.001	.0372383	.149713
nagencph	-6.994537	.6091561	-11.48	0.000	-8.188471	-5.800603
hempr	2.479602	.1591608	15.58	0.000	2.16765	2.791554
size	-.1777728	.0602048	-2.95	0.003	-.295773	-.0597727
_cons	75.42751	1.590008	47.44	0.000	72.31113	78.5439

Estimação da Equação 1.5 (*dummies* de região e IF não reportados)

i.reg	_Ireg_1-54	(naturally coded; _Ireg_1 omitted)				
i.idif	_Iidif_1-34	(naturally coded; _Iidif_1 omitted)				
Source	SS	df	MS	Number of obs = 143636		
Model	1163197.66	96	12116.6423	F(96,143539) = 154.15		
Residual	11282841.1143539	78	6047071	Prob > F = 0.0000		
Total	12446038.7143635	86	6504592	R-squared = 0.0935		
				Adj R-squared = 0.0929		
				Root MSE = 8.8659		
prazo	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
subj1	-.3267257	.0101522	-32.18	0.000	-.3466239	-.3068275
subp1	.0076611	.0284962	0.27	0.788	-.048191	.0635131
relac	-.1521426	.0076624	-19.86	0.000	-.1671608	-.1371245
relac2	.1021863	.0144903	7.05	0.000	.0737856	.1305871
risco	1.402841	.0657973	21.32	0.000	1.27388	1.531803
lrenda	-.0295239	.0322137	-0.92	0.359	-.0926622	.0336144
totif	-.0456366	.0259711	-1.76	0.079	-.0965395	.0052662
nagencph	-3.593746	.5513729	-6.52	0.000	-4.674426	-2.513066
hempr	.2720087	.1440632	1.89	0.059	-.0103524	.5543697
size	-.1486865	.0544939	-2.73	0.006	-.2554934	-.0418796
_cons	35.06232	1.439183	24.36	0.000	32.24155	37.88309

APÊNDICE 1.C

Com as estimativas do **Modelo 9**, calculamos os valores das elasticidades para cada observação e computamos o seu valor médio e o seu respectivo desvio-padrão. Na **Tabela A - 1.2** apresentamos o valor para a amostra total e na **Tabela A - 1.3**, para as subdivisões da amostra.

Tabela A - 1.2 – Elasticidade da demanda por financiamento de veículos – Modelo 9

taxa de juros		prazo	
-0,416	(1,646)	+0,393	(0,506)

Coefficientes não-significativos. Desvio-padrão da amostra entre parênteses.

Tabela A - 1.3 – Elasticidades para o Modelo 9 as diversas subamostras

Divisão por:	Valor da prestação		Renda		Valor do empréstimo		Comprometimento da renda	
Variável	1/3 inferior	1/3 superior	1/3 inferior	1/3 superior	1/2 inferior	1/2 superior	<= 0,25	>0,25
<i>jur</i>	-0,267* (0,050)	-0,256 (1,264)	-0,441* (0,152)	-0,383 (2,245)	-0,108 (0,287)	-0,309 (0,966)	-0,377* (0,094)	-0,444 (2,675)
<i>prazo</i>	0,963* (0,359)	0,972* (0,354)	0,529 (0,495)	0,242 (0,529)	0,054 (0,043)	0,348 (0,604)	0,316 (0,276)	0,924* (0,318)
<i>Nº obs:</i>	47.879	47.879	47.879	47.879	71.818	71.818	97.746	45.890

Em azul, o subgrupo mais provável de sofrer restrição. Desvio-padrão da amostra entre parênteses.

* Coeficiente significativo a 5%.

Capítulo 2

2 Um Estudo sobre a Relação entre Taxas de Juros, Garantias e Tempo de Relacionamento nos Contratos de Crédito

2.1 INTRODUÇÃO

Este trabalho empírico utiliza os dados do SCR – Sistema de Informações de Crédito do Banco Central do Brasil, na modalidade Capital de Giro, e tem dois objetivos. O primeiro é testar a relação existente entre taxas de juros e garantias; o segundo é verificar qual o impacto do tempo de relacionamento entre o tomador e o banco nesses contratos de crédito.

Os bancos são instituições cujas atividades principais são receber depósitos e conceder empréstimos, sendo que ambas estão inter-relacionadas. Quando um banco concede empréstimo a um tomador, existe o risco da inadimplência¹⁷. Se o banco ajusta o risco cobrando uma taxa de juros mais alta, ele vai atrair os clientes mais arriscados (seleção adversa) e com maior probabilidade de levar adiante projetos de maior risco (risco moral). A princípio, tanto a existência de garantias contratuais quanto um maior tempo de relacionamento podem contribuir para diminuir este problema, reduzindo portanto a taxa de juros. No entanto, os trabalhos teóricos e empíricos não são unânimes em estabelecer essa relação entre tais variáveis.

Os principais resultados acerca de como garantias e risco de crédito se associam são resumidos por Berger & Udell (1990) em duas abordagens: o paradigma do risco observado

¹⁷ Além do risco da concessão do empréstimo, o banco também está exposto ao risco de taxa de juros e risco de liquidez, vide Freixas e Rochet (1997).

(*sorting-by-observed-risk paradigm*) e o paradigma da informação privada (*sorting-by-private-information paradigm*).

No primeiro caso, assume-se que os bancos observam o risco dos clientes e os tomadores arriscados são solicitados a apresentar garantias enquanto os seguros não. Os bancos devem possuir informações que lhes permitam distinguir os tomadores com relação ao risco, mas nada afirma sobre o risco do empréstimo. E um empréstimo para um cliente de alto risco será condicionado à existência de garantias.

No paradigma da informação privada, devido à existência de informação assimétrica, não é possível aos bancos distinguir os tomadores com base no seu risco. Assume-se que os bancos decidem sobre a garantia e a taxa de juros simultaneamente. Neste caso, os tomadores de alto risco podem ser identificados porque selecionam os contratos com pouca garantia e taxas de juros altas (Bester, 1985). Existe, portanto, uma associação negativa entre essas variáveis.

O impacto do tempo de relacionamento nos termos do contrato também tem predições opostas na literatura. Por um lado, um longo relacionamento pode resultar em contratos vantajosos para o tomador. Isto ocorre pois os bancos podem taxar e subsidiar os tomadores ao longo do tempo, e desta forma podem reduzir o uso da garantia e/ou da taxa de juros de modo a diminuir os problemas de risco moral (Boot e Thakor, 1994). Por outro lado, um relacionamento duradouro pode se mostrar prejudicial ao tomador. Ao longo desse processo, o banco acumula mais informações acerca do cliente que os demais competidores. Essa assimetria faz com que o banco seja monopolista das informações permitindo a ele cobrar maior taxa de juros e exigir maiores garantias dos clientes mais antigos (Sharpe, 1990).

Estimamos a associação entre taxa de juros e garantias controlando pelo tempo de relacionamento, pela maturidade e por outras variáveis, utilizando diversas técnicas econométricas. A taxa de juros é o $\ln$ da taxa de juros em % a.a. e a garantia é uma variável binária igual a 1 se o contrato apresenta algum tipo de garantia não fidejussória e zero caso contrário. Como não estamos certos do processo que determina ambas variáveis, estimamos duas equações, um *probit* com a garantia como variável dependente e a taxa de juros como variável explicativa, e por OLS e 2SLS invertendo o papel dessas variáveis. Propomos

também outros métodos adicionais para a estimação da equação da taxa de juros: o *matching estimator*, o efeito médio do tratamento com endogeneidade e o *switching regression model*.

Todas as técnicas utilizadas indicam que a taxa de juros e garantias apresentam relação negativa nos empréstimos de Capital de Giro. Este resultado é compatível com *sorting-by-private-information paradigm*: como os bancos não conseguem distinguir os tomadores pelo risco, ofertam contratos com garantia e baixa taxa de juros e sem garantia e alta taxa de juros. No equilíbrio separador, cada tipo escolhe o contrato que lhe foi designado.

Tanto na equação de garantia quanto na equação de taxa de juros, o tempo de relacionamento impacta negativamente. Dessa forma, obtemos evidências de que um cliente antigo da instituição pode ser beneficiado com contratos mais vantajosos. A conclusão geral deste capítulo é que existe indícios de que no mercado de crédito brasileiro (na modalidade capital de giro) está presente a assimetria informacional e um tempo de relacionamento mais longo ajuda a diminuir os seus efeitos.

2.2 HIPÓTESES EMPÍRICAS E REVISÃO DA LITERATURA

2.2.1 Relação entre taxa de juros e garantia

Existem diversas razões pelas quais taxa de juros e garantia possam estar relacionadas negativamente. Tudo o mais constante, um empréstimo com garantias tende a reduzir as perdas dos bancos em caso de default. Outro ponto em favor desta hipótese é que o problema de risco moral que advém após a concessão do crédito é menor quando o empréstimo possui garantias. As decisões das firmas tendem a ser menos arriscadas, pois em caso de inadimplência a garantia pode ser executada. Neste caso a garantia é um mecanismo que faz com que o contrato seja exequível e os interesses dos tomadores estejam alinhados aos dos bancos. Finalmente, as garantias ajudam a mitigar os problemas de seleção adversa, pois constituem um importante instrumento sinalizador da qualidade do tomador. Se os bons pagadores estão dispostos a ofertar garantia em troca de uma taxa de juros mais baixa (que os maus pagadores), então a garantia é uma fonte valiosa de informação, indicando a qualidade de quem está tomando o crédito.

A base teórica para essa linha de argumento é fornecida por Chan e Kanatas (1985) e Bester (1985). Utilizando os dados do SCR do Banco Central do Brasil (*cross-section* de dez/2003), Rodrigues *et al.* (2004) encontram evidências empíricas que indicam que este resultado é válido para o Brasil. Os dados dos contratos belgas para pequenas empresas também fornecem indícios em favor dessa tese (Degryse e Van Cayseele, 2000).

No entanto, nem todos os resultados teóricos e empíricos são favoráveis à situação acima descrita e a garantia pode estar associada ao maior risco de crédito. Manove *et al.* (2001) elaboram um modelo no qual os bancos têm menos incentivos a fazer a seleção adequada dos projetos das firmas (*screening*) se os contratos estão protegidos pela garantia (*lazy banks*). Ainda, podem existir empresários demasiado otimistas com o projeto e que estão dispostos a prover todas as garantias necessárias para obter o financiamento. Nesse caso, mais garantia implicaria maiores perdas esperadas pelos credores. Para Boot *et al.* (1991), na presença de risco moral, se os credores conseguem distinguir a qualidade do tomador, os mais arriscados terão um contrato com garantias e os menos arriscados sem. Neste modelo, a garantia serve para que o tomador arriscado depreenda mais esforço no projeto, uma vez que reduz a

probabilidade de default e conseqüentemente a perda da garantia. Alguns trabalhos empíricos que corroboram esta tese são apresentados a seguir. Berger e Udell (1990) encontram evidências de que existe uma relação positiva entre garantia e risco (medida como prêmio de risco) para os dados americanos. Resultado análogo é obtido por Jiménez e Saurina (2004) usando os dados da *Central de Información de Riesgos* do Banco Central Espanhol. Os autores investigam qual o impacto da garantia na probabilidade dos empréstimos bancários sofrerem default e encontram evidências de que empréstimos com garantia têm maiores chance de inadimplência.

Este debate pode ser resumido em duas hipóteses relacionadas a situações em que o risco pode ou não ser observado pelo banco (Berger e Udell, 1990): **1)** se o banco consegue distinguir clientes arriscados, então a concessão do empréstimo estará atrelada à apresentação de garantias (tomadores arriscados e empréstimos arriscados ou não); e **2)** por outro lado, se a informação sobre o risco do cliente é privada, então a garantia é apresentada pelos tomadores menos arriscados (tomadores e empréstimos pouco arriscados). O primeiro caso é conhecido na literatura como paradigma do risco observado (*sorting-by-observed-risk paradigm*) e o segundo, paradigma da informação privada (*sorting-by-private-information paradigm*).

Neste contexto, desejamos verificar se no Brasil taxas de juros e garantia apresentam relação positiva ou negativa (ou ainda, inexistente). Nossa primeira predição a ser testada empiricamente é:

Hipótese 1: *Se os bancos são capazes de distinguir os clientes pelo risco, a taxa de juros e a garantia são variáveis complementares num contrato de crédito. Por outro lado, na presença de assimetria informacional, a garantia serve para sinalizar a qualidade do tomador e, portanto, essas variáveis são substitutas no contrato de crédito.*

Existem ainda outros resultados referentes a essa literatura de garantia e risco. Em trabalho recente, Ono e Uesugi (2008) investigam os determinantes do uso da garantia (garantia real e pessoal) nos contratos de crédito de pequenas e médias empresas no Japão. Os autores encontram evidências de que o risco do tomador não influi na probabilidade do contrato possuir garantia.

Weill e Godlewski (2006), por sua vez, conduzem seus estudos com dados de empréstimos bancários de 43 países e verificam qual é a relação entre a garantia e o prêmio de risco. Os autores incluem, como *proxy* da assimetria informacional, medidas de opacidade de informações financeiras (qualidade das informações), mecanismos de compartilhamento de informações, *trust* e medidas de desenvolvimento econômico e financeiro. O resultado encontrado é que a assimetria influi na relação entre garantia e prêmio de risco: um menor grau de assimetria informacional aumenta a associação positiva entre ambas variáveis suportando a tese de que o risco do tomador é observado. Também existem evidências acerca da hipótese de seleção adversa, pois um maior grau de assimetria informacional reduz essa relação positiva. Segundo os autores, ambas as hipóteses podem ser validadas empiricamente, dependendo do grau de assimetria presente no país.

2.2.2 Tempo de relacionamento entre o tomador e o banco

Outro ponto no qual as predições da literatura de crédito são divergentes refere-se ao tempo de relacionamento do tomador com a IF. Um longo tempo de relacionamento pode beneficiar o tomador, pois permite que o banco o conheça melhor, diminuindo os problemas decorrentes da assimetria informacional (Boot e Thakor, 1994). A predição empírica é que tanto a taxa de juros como a garantia diminuem com o tempo de relacionamento. Por outro lado, um relacionamento mais duradouro pode ser prejudicial ao tomador na medida em que os bancos podem usar essa vantagem informacional para exercer o seu poder de monopólio (Sharpe, 1990). Empiricamente, espera-se que tanto a taxa de juros quanto a garantia aumentem com o tempo de relacionamento.

Os resultados encontrados na literatura empírica são apresentados na seqüência. Berger e Udell (1995) trabalham com as linhas de crédito para pequenas empresas americanas e desejam testar se um relacionamento mais longo diminui a assimetria informacional e, por conseguinte, como os termos dos contratos são ajustados. Os resultados sugerem que o tempo de relacionamento é informativo acerca da qualidade do tomador, reduzindo a taxa de juros e a probabilidade do contrato ter garantia. Jiménez, Salas e Saurina (2004) utilizando dados de contratos espanhóis encontram evidências de que o relacionamento (tempo e escopo) diminui a probabilidade de o empréstimo apresentar garantia.

Por outro lado, os resultados de Ono e Uesugi (2008) utilizando dados de pequenas e médias empresas japonesas sugerem que quanto maior o tempo de relacionamento com o principal credor, mais provável que o contrato apresente garantia. Degryse e Van Cayseele (2000), utilizando dados de empresas de pequeno porte belgas concluem que o tempo de relacionamento influi positivamente e o escopo (quantidade de produtos e serviços comprados do banco) negativamente na taxa de juros, sendo este último o mais importante. Resultado oposto é encontrado para a garantia.

Dada as considerações teóricas e empíricas acerca do tempo de relacionamento entre o tomador e o banco, a segunda hipótese empírica que queremos testar para o caso brasileiro é:

Hipótese 2: *A taxa de juros e a garantia diminuem com o tempo de relacionamento se a mesma reduz a assimetria informacional. Por outro lado, elas aumentam se o efeito do hold-up predominar.*

2.3 ESTIMAÇÃO

Nesta seção iremos apresentar as diversas técnicas econométricas utilizadas para testar as duas hipóteses. Inicialmente, estimamos a equação da garantia tendo como variáveis explicativas o $\ln$ da taxa de juros, o tempo de relacionamento e outros controles. Como a garantia é uma variável binária e a taxa de juros pode ser endógena, propomos a estimação por *probit* com e sem endogeneidade.

A segunda equação a ser estimada é do $\ln$ da taxa de juros tendo como variáveis explicativas a garantia (possivelmente endógena), o relacionamento e outros controles. Inicialmente estimamos por OLS – *Ordinary Least Squares* e 2SLS – *Two-stage Least Squares*. Na sequência, utilizamos as técnicas comumente empregadas na avaliação de programas de treinamento e tratamento de saúde e propomos a técnica do *matching estimator* e do *average treatment effects*. Neste caso, os contratos com e sem garantia representam o tratamento e o não-tratamento, respectivamente. A última técnica empregada refere-se ao *endogenous switching regression model* e, neste caso, a hipótese subjacente é que os contratos com e sem garantia podem ser representados por dois regimes distintos. Os instrumentos utilizados são apresentados no final desta seção.

2.3.1 Equação da garantia

Inicialmente, queremos verificar o impacto da taxa de juros na probabilidade de um empréstimo apresentar garantias. A equação que queremos modelar é:

Equação 2.1:
$$gar^* = z_1\delta_1 + \alpha_1 l_{juros} + u_1$$

$$gar = 1[gar^* > 0]$$
Onde $1[.]$ é uma função indicadora.

A variável dependente gar é igual a um se o empréstimo possui garantias não fidejussórias e zero caso contrário e por isso estimamos a **Equação 2.1** como um *probit*. Se a taxa de juros for endógena, então ela deve ser instrumentalizada por z_2 :

Equação 2.2: $l_{juros} = z_1\delta_{21} + z_2\delta_{22} + v_2$

z_1 é o vetor de variáveis de controle suposto exógeno; δ_1 , δ_{21} , δ_{22} e α_1 são os (vetores de) coeficientes a serem estimados; (u_1, v_2) têm média zero, distribuição bivariada normal e é independente de z .

A taxa de juros será endógena se u_1 e v_2 são correlacionados. A **Equação 2.1** juntamente com a **Equação 2.2** é denominada de **probit com variável explicativa endógena contínua** (Wooldridge (2002), p. 472).

2.3.2 Equação da taxa de juros

Para estimar os efeitos da garantia na taxa de juros, utilizamos o seguinte modelo:

Equação 2.3: $l_{juros} = X\beta + \theta_1 gar + \varepsilon$

Onde X é o vetor de variáveis exógenas e ε o erro.

Se a variável gar é exógena, podemos estimar os coeficientes da **Equação 2.3** por OLS (Mínimos Quadrados Ordinários). Por outro lado, se $cov(gar, \varepsilon) \neq 0$, então utilizamos o método das variáveis instrumentais (IV) e, para isso, necessitamos de instrumento(s) Z tal(is) que: 1) Z não é correlacionado com ε ; e 2) Z é (parcialmente) correlacionado com a variável gar . Conforme Wooldridge (2002), p. 84, tanto a endógena quanto os instrumentos podem ser variáveis contínuas ou discretas¹⁸.

¹⁸ A projeção linear da variável endógena nas variáveis exógenas e instrumentos está definida quando os segundos momentos de todas as variáveis são finitos.

2.3.2.1 Matching Estimator

Uma segunda forma para estimar a **Equação 2.3** é calcular o efeito médio das garantias reais sobre a taxa de juros dos contratos de crédito (*average treatment effect* – ATE¹⁹). A técnica utilizada é o *matching estimator*, cuja idéia principal é a seguinte. Seja $Y(1)$ a taxa de juros quando o empréstimo possui garantias e $Y(0)$ a taxa de juros quando o empréstimo não possui garantias. Suponha que observamos $Y(1)$ e queremos estimar $Y(0)$. Se as decisões de apresentar (ou solicitar) garantias são aleatórias para contratos com as mesmas características, então a média das taxas de juros de contratos similares sem garantias podem ser utilizadas para estimar $Y(0)$. Neste trabalho, implementamos o estimador com correção de viés²⁰ proposto por Abadie e Imbens (2007) e Abadie *et al.* (2004, para o comando *nnmatch* do Stata). Esses autores definem:

$$\text{Equação 2.4:} \quad Y = Y(W) = \begin{cases} Y(0) & \text{se } W = 0 \\ Y(1) & \text{se } W = 1 \end{cases}$$

E, para todo x no suporte de X :

1. W é independente de $(Y(0), Y(1))$ condicional em $X = x$, ou seja, para indivíduos similares, a escolha do tratamento é aleatória; e
2. $c < \Pr(W = 1 / X = x) < 1 - c$, para algum $c > 0$.

Onde X são as covariadas, $W \in \{0,1\}$ é uma variável binária que indica o tratamento recebido e Y é a realização.

Para cada indivíduo i , $i = 1, \dots, N$, considere X_i as suas covariadas observadas tal que

$\|x\|_V = (x'Vx)^{1/2}$ é a norma do vetor com matriz positiva definida V . A distância entre os

¹⁹ Esta técnica é bastante utilizada na literatura de avaliação de programas.

²⁰ O estimador de *matching* simples será viesado em amostras finitas quando mais de uma variável contínua for utilizada para fazer o *matching*. Os autores propõem um estimador (*bias-corrected matching estimator*) que remove assintoticamente esse viés condicional.

vetores x e z é definida como $\|z - x\|_V$, onde z são os valores das covariadas de um potencial *matching* para a observação i . Seja $d_M(i)$ a distância²¹ entre X_i e o M -ésimo *match*²² que possui tratamento oposto. O conjunto dos M primeiros índices que fazem *matching* com a observação i é denominado $\mathfrak{S}_M(i)$:

$$\mathfrak{S}_M(i) = \{l = 1, \dots, N / W_l = 1 - W_i, \|X_l - X_i\|_V \leq d_M(i)\}$$

O número de elementos de $\mathfrak{S}_M(i)$ é denominado $\#\mathfrak{S}_M(i)$. Finalmente, seja $K_M(i)$ o número de vezes que a observação i é utilizada para fazer o *matching* para todas as l observações do grupo de tratamento oposto, ponderada pelo número total de *matches* para a observação l :

$$K_M(i) = \sum_{l=1}^N 1\{i \in \mathfrak{S}_M(l)\} \frac{1}{\#\mathfrak{S}_M(l)}$$

As potenciais realizações são estimadas como:

$$Y_i^{\hat{}}(0) = \begin{cases} Y_i & \text{se } W_i = 0 \\ \frac{1}{\#\mathfrak{S}_{M,i}} \sum_{l \in \mathfrak{S}_M(i)} Y_l & \text{se } W_i = 1 \end{cases}$$

e

$$Y_i^{\hat{}}(1) = \begin{cases} \frac{1}{\#\mathfrak{S}_{M,i}} \sum_{l \in \mathfrak{S}_M(i)} Y_l & \text{se } W_i = 0 \\ Y_i & \text{se } W_i = 1 \end{cases}$$

²¹ Formalmente, $d_M(i) > 0$ é um número real que satisfaz as seguintes condições:

$\sum_{l: W_l = 1 - W_i} 1\{\|X_l - X_i\|_V < d_M(i)\} < M$ e $\sum_{l: W_l = 1 - W_i} 1\{\|X_l - X_i\|_V \leq d_M(i)\} \geq M$, onde $1\{\cdot\}$ é uma função indicadora

igual a 1 se a expressão entre chaves é verdadeira e zero caso contrário.

²² Dizemos que uma observação i , que faz parte do grupo de tratamento, faz *matching* com a observação l , que faz parte do grupo de controle, se as suas covariadas são próximas.

Ou seja, para cada realização i observada, calculamos a média das l realizações com as mesmas covariadas que i , mas que pertencem ao grupo de tratamento oposto. Utilizando estas estimativas, o *simple matching estimator* (ATE) é computado como:

$$\hat{\tau}_M^{sm} = \frac{1}{N} \sum_{i=1}^N \{ \hat{Y}_{i,1} - \hat{Y}_{i,0} \}$$

O efeito médio do tratamento para os tratados é (ATT):

$$\hat{\tau}_M^{sm,t} = \frac{1}{N_1} \sum_{i:W_i=1}^N \{ Y_i - \hat{Y}_{i,0} \}$$

O efeito médio do tratamento para grupo de controle é (ATC):

$$\hat{\tau}_M^{sm,c} = \frac{1}{N_0} \sum_{i:W_i=0}^N \{ \hat{Y}_{i,1} - Y_i \}$$

Onde $N_1 = \sum_{i=1}^N W_i$ é o número de tratados e $N_0 = \sum_{i=1}^N (1 - W_i)$ é o número de controles.

O *simple matching estimator* será viesado em amostras finitas quando o *matching* não for exato. Para corrigir este problema, Abadie e Imbens (2007) propõem o *bias-corrected matching estimator*. O ajuste é baseado na estimativa de duas regressões, $\hat{\psi}_w(x) = E\{Y(w) / X = x\}$, para $w = 0, 1$. Essas regressões são aproximadas por uma função linear e estimadas usando mínimos quadrados nas observações que fazem o *matching*. As previsões das potenciais realizações ausentes (*missing*) para o estimador com correção de viés são dadas por:

$$Y_i \quad \text{se } W_i = 0$$

$$Y_i^{\sim}(0) = \frac{1}{\#\mathfrak{S}_{M,i}} \sum_{l \in \mathfrak{S}_M(i)} \{Y_l + \hat{\psi}_0(X_i) - \hat{\psi}_0(X_l)\} \quad \text{se } W_i = 1$$

E

$$Y_i^{\sim}(1) = \frac{1}{\#\mathfrak{S}_{M,i}} \sum_{l \in \mathfrak{S}_M(i)} \{Y_l + \hat{\psi}_1(X_i) - \hat{\psi}_1(X_l)\} \quad \text{se } W_i = 0$$

$$Y_i \quad \text{se } W_i = 1$$

E os correspondentes *bias-corrected matching estimators* para o ATE, ATT e ATC são respectivamente iguais a:

$$\hat{\tau}_M^{bcm} = \frac{1}{N} \sum_{i=1}^N \{\tilde{Y}_{i,1} - \tilde{Y}_{i,0}\},$$

Equação 2.5:

$$\hat{\tau}_M^{bcm,t} = \frac{1}{N_1} \sum_{i:W_i=1}^N \{Y_i - \tilde{Y}_{i,0}\}, \text{ e}$$

$$\hat{\tau}_M^{bcm,c} = \frac{1}{N_0} \sum_{i:W_i=0}^N \{\tilde{Y}_{i,1} - Y_i\}.$$

2.3.2.2 Average Treatment Effect

Pode-se também estimar o ATE considerando-se a endogeneidade. A teoria aqui exposta segue Wooldridge (2002, capítulo 18). Seja W uma variável binária, tal que $W = 1$ denota o recebimento do tratamento e $W = 0$, caso contrário, sendo que o tratamento significa o empréstimo apresentar garantias reais.

Sejam Y_1 e Y_0 a taxa de juros do contrato quando o mesmo recebe e não recebe o tratamento. Se pudéssemos observar os termos dos contratos nas situações com e sem garantias, então o

efeito médio do tratamento seria diretamente calculável e igual a $ATE = E(Y_1 - Y_0)$ e o efeito médio do tratamento nos tratados seria $ATT = E(Y_1 - Y_0 / W = 1)$. Na prática quase nunca observamos Y_1 e Y_0 para um mesmo contrato e o problema que defrontamos é similar ao de ausência de dados (*missing data*)²³. Dado W , o que observamos é:

Equação 2.6:
$$Y = (1 - W)Y_0 + WY_1 = Y_0 + W(Y_1 - Y_0)$$

Defina $Y_0 = \mu_0 + \nu_0$ e $Y_1 = \mu_1 + \nu_1$, onde $\mu_g = E(Y_g)$ e $E(\nu_g) = 0$, $g = 0, 1$. Substituindo na **Equação 2.6**, temos:

Equação 2.7:
$$Y = \mu_0 + (\mu_1 - \mu_0)W + \nu_0 + W(\nu_1 - \nu_0)$$

Se o tratamento é endógeno, então devemos instrumentalizá-lo. Adicionalmente, assumimos:

- Hipótese A:**
- a) Os termos estocásticos de Y_1 e Y_0 são iguais ($\nu_1 = \nu_0$);
 - b) $L(\nu_0 / X, Z) = L(\nu_0 / X)$; e
 - c) $L(W / X, Z) \neq L(W / X)$.

Onde $L(\cdot)$ é a projeção linear que contém o termo do intercepto, X é o vetor de covariadas e Z é o vetor de instrumentos.

Sob os itens **a)** e **b)** da **Hipótese A**, podemos reescrever a **Equação 2.7**²⁴ como:

Equação 2.8:
$$Y = \omega_0 + \tau W + X\beta + u_0$$

Onde $\tau = ATE$ e $u_0 = \nu_0 - L(\nu_0 / X, Z)$. Esta equação pode ser consistentemente estimada, conforme já mencionado, por 2SLS (Mínimos Quadrados de Dois Estágios). No entanto, com hipóteses mais fortes, podemos encontrar um estimador IV mais eficiente:

²³ Vide Wooldridge (2002), pg. 604.

²⁴ Esta equação é conhecida como **modelo de variável endógena dummy** (Heckman, 1978).

Hipótese A’: a) Os termos estocásticos de Y_1 e Y_0 são iguais ($\nu_1 = \nu_0$);

b) $E(\nu_0 / X, Z) = L(\nu_0 / X)$; e

c) $P(W = 1 / X, Z) \neq P(W = 1 / X)$ e $P(W = 1 / X, Z) = G(X, Z; \gamma)^{25}$ é uma forma paramétrica conhecida (*logit* ou *probit*); e

d) $Var(\nu_0 / X, Z) = \sigma_0^2$

Sob a **Hipótese A’**, uma forma para estimar a **Equação 2.8** ²⁶: 1) Estima-se $P(W = 1 / X, Z) = G(X, Z; \gamma)$ por máxima verossimilhança, obtendo as probabilidades ajustadas $\hat{G}_i$; e 2) Estima-se a **Equação 2.8** por IV usando 1, $\hat{G}_i$ e X_i como instrumento para W_i .

2.3.2.3 Switching Regression Model

O *switching regression model with endogenous switching* permite que o termo de intercepto e os coeficientes das demais variáveis sejam diferentes entre as operações com e sem garantia. Neste caso, estamos interessados na seguinte especificação (vide Maddala, 1983, capítulo 8):

$$\begin{aligned} \text{Equação 2.9:} \quad & \text{Regime 1: } ljur_{1i} = X_{1i}\beta_1 + u_{1i} & \text{se } I_i = 1 \\ & \text{Regime 2: } ljur_{2i} = X_{2i}\beta_2 + u_{2i} & \text{se } I_i = 0 \end{aligned}$$

$$I_i = 1 \text{ se } \gamma Z_i + u_i < 0$$

$$I_i = 0 \text{ caso contrário.}$$

²⁵ Na **Equação 2.8**, τ e β podem ser identificados mesmo que não haja variáveis exógenas extras excluídas de X , pois $G(X; \gamma)$ é uma função não-linear de X .

²⁶ Conforme Wooldridge (2002) p.623, este procedimento tem a propriedade de ser robusto. Como se utiliza $\hat{G}_i$ como instrumento para W_i , o modelo para $P(W = 1 / X, Z)$ não precisa ser corretamente especificado. Geralmente, o que se necessita é que a projeção de W em $[X, G(X, Z; \gamma^*)]$ dependa de $G(X, Z; \gamma^*)$, onde γ^* denota o plim do estimador de máxima verossimilhança quando o modelo não está bem especificado.

Onde I_i é uma variável binária que determina se o contrato de crédito i tem ou não garantias. $Ljur_{1i}$ e $ljur_{2i}$ são o $\ln$ da taxa de juros do contrato i com e sem garantias, respectivamente, e são as variáveis dependentes da equação contínua. X_{1i} e X_{2i} são os vetores de variáveis exógenas. Z_i é o vetor de características que influenciam na decisão de o empréstimo possuir ou não garantias. β_1 , β_2 e γ são os vetores de parâmetros a serem estimados. Assume-se que u_{1i} , u_{2i} e u_i são os termos de erro que seguem uma distribuição normal trivariada com média zero e matriz de variância-covariância Ω :

$$\Omega = \begin{bmatrix} \sigma_u^2 & \sigma_{1u} & \sigma_{2u} \\ & \sigma_1^2 & \sigma_{12} \\ & & \sigma_2^2 \end{bmatrix}$$

Adota-se a hipótese de que a variância do erro da equação de seleção é igual a um ($\sigma_u^2 = 1$). σ_1^2 e σ_2^2 são as variâncias do erro das equações contínuas. σ_{1u} é a covariância entre u_{1i} e u_i e σ_{2u} é a covariância entre u_{2i} e u_i . Se $\sigma_{1u} = \sigma_{2u} = 0$ então o *switching* deve ser exógeno. σ_{12} não está definido, pois nunca observamos $ljur_{1i}$ e $ljur_{2i}$ simultaneamente.

A esperança não-condicional do $\ln$ da taxa de juros é:

Equação 2.10: $E(ljur_{1i}) = X_{1i}\beta_1$
 $E(ljur_{2i}) = X_{2i}\beta_1$

A esperança condicional do $\ln$ da taxa de juros é:

Equação 2.11: $E(ljur_{1i} / I_i = 1) = X_{1i}\beta_1 - \sigma_1\rho_1 \frac{\phi(\gamma Z_i)}{1 - \Phi(\gamma Z_i)}$
 $E(ljur_{2i} / I_i = 0) = X_{2i}\beta_2 + \sigma_2\rho_2 \frac{\phi(\gamma Z_i)}{\Phi(\gamma Z_i)}$

Onde $\rho_1 = \frac{\sigma_{1u}}{\sigma_1\sigma_u}$ e $\rho_2 = \frac{\sigma_{2u}}{\sigma_2\sigma_u}$ são os coeficientes de correlação entre os erros (u_{1i}, u_i) e (u_{2i}, u_i) , respectivamente. $\phi(\cdot)$ é a distribuição de densidade normal e $\Phi(\cdot)$ é a função de densidade cumulativa normal.

Pela **Equação 2.11**, observa-se que, neste modelo, a média condicional é influenciada por ρ_1 e ρ_2 . Se $\rho_1 > 0$, então conclui-se que o *ln* da taxa de juros de um contrato com garantias é menor do que em um contrato aleatório. Se $\rho_1 < 0$, então o *ln* da taxa de juros de um contrato com garantias é maior do que em um contrato aleatório. Se $\rho_1 = 0$, então $\sigma_{1u} = 0$ e a mudança de regime deve ser exógena. Resultado análogo é obtido para o coeficiente ρ_2 .

O sistema representado pela **Equação 2.9** está identificado mesmo que $X = Z$, devido à não-linearidade da forma funcional da equação binária. No entanto, estimadores mais confiáveis podem ser obtidos incluindo restrições de identificação. Neste caso, necessitamos de variáveis que influenciam a seleção do regime, mas que não influenciam a taxa de juros da operação. As variáveis que utilizamos para este fim são um índice de eficiência do judiciário e *dummies* para a região onde o empréstimo foi concedido²⁷. A hipótese é que essas variáveis não influenciam a taxa de juros, dada a escolha do regime.

Os resultados apresentados na parte empírica partem do pressuposto de que a existência ou não da garantia é endógena na determinação da taxa de juros, cujos erros u_{1i} , u_{2i} e u_i seguem uma distribuição normal trivariada $(0, \Omega)$. A implementação dessa técnica é feita por FIML (*Full-Information Maximum Likelihood*) para ajustar simultaneamente a parte binária e contínua do modelo²⁸.

2.3.3 Instrumentos

Alguns autores não tratam desse potencial problema da endogeneidade da garantia na equação da taxa de juros. Berger e Udell (1990, 1995) e Degryse e Van Cayseele (2000) partem do

²⁷ Mais informações acerca destas variáveis serão fornecidas ao longo do capítulo.

²⁸ A estimação é feita através do comando *movestay* do Stata, desenvolvido por Lokshin e Sajaia (2004).

pressuposto de que garantia e taxa de juros são determinados de modo sequencial, primeiro decidindo-se acerca da garantia e depois a taxa de juros.

Num contrato de empréstimo para Capital de Giro, é provável que a garantia e a taxa de juros sejam determinadas conjuntamente, quer o banco consiga ou não distinguir os clientes pelo risco. Em qualquer situação, a existência ou não de garantias deve estar relacionada com as características da empresa que toma o crédito.

Uma grande vantagem de se utilizar os dados do SCR é que obtemos os preços (taxas de juros) de cada contrato e não a sua média. E por isso, se consideramos que os consumidores são tomadores de preços, a endogeneidade que pode advir do erro de medida é menos preocupante na estimação. O segundo benefício de se trabalhar com dados individualizados é que conseguimos controlar parcialmente a heterogeneidade não observada. O SCR traz algumas características individuais tais como o tempo de relacionamento, o número de IFs com a qual o tomador está relacionado e o seu porte.

No presente caso, na equação da garantia, as variáveis instrumentais para a taxa de juros têm as mesmas motivações que as utilizadas na **subseção 1.4.2**, p.26 do **Capítulo 1**. Utilizamos as *dummies* de IFs como instrumentos deslocadores de oferta, pois elas captam todas as diferenças relativas a custos existentes entre as instituições financeiras. O índice *Herfindahl* de depósitos (*hdep*) mede o grau de concentração e o volume de crédito dividido pelo PIB do município (*pot*) mede o potencial do mercado. Ambas as variáveis foram calculadas para o município em que a agência que efetuou a transação está localizado.

Na equação da taxa de juros, uma variável que seja correlacionada com a garantia e não seja correlacionada com o erro é um candidato a instrumento. Em caso de inadimplência, o primeiro modo para recuperar o saldo devedor é a geração de caixa do cliente. As garantias constituem a segunda alternativa de recuperação do empréstimo e podem incidir sobre coisas próprias ou alheias. As garantias podem ser fidejussórias (garantias pessoais) ou não fidejussórias (garantias reais)²⁹. Nas garantias reais, existe a destinação de um bem (móvel ou

²⁹ As garantias não fidejussórias podem ser: a alienação fiduciária, o penhor, a hipoteca ou a anticrese. O aval e a fiança são tipos de garantias fidejussórias.

imóvel) que em caso de inadimplência será objeto de execução para liquidação da dívida (após decisão judicial) (vide Ortolani, 2000, p.53).

A garantia real pode ser um título de crédito, mercadorias e bens móveis ou imóveis. A princípio, não existe objeção para que uma delas seja preferida em relação a outra. No entanto, a maior ou menor facilidade de avaliação pode fazer com que os títulos de crédito sejam mais utilizados como garantia. Nossa hipótese é a de que as diversas formas de garantia guardam alguma substituição, ainda que imperfeita, e, portanto, devem estar correlacionadas entre si. A variável que utilizamos para representar as diversas formas de garantia é *cap*, que é igual ao valor médio de um imóvel residencial em um município. O valor do aluguel dos imóveis deve ser altamente correlacionado com o valor dos imóveis, que por sua vez servirá de garantia no contrato de crédito. Portanto, a variável *cap*³⁰ deve estar correlacionada com a garantia, influenciando a taxa de juros indiretamente através dela. A vantagem da utilização deste instrumento é que o mesmo é calculado para cada município, apresentando grande variabilidade.

É provável que o sucesso na recuperação das garantias dependa da eficiência do judiciário. De pouco adianta uma garantia que valha várias vezes o valor do empréstimo se o direito para a sua execução ocorre muitos anos mais tarde. O problema se agrava nos casos em que o bem se deprecia ao longo do tempo, como é o caso dos veículos, por exemplo. Apesar de o caminho legal a ser percorrido por um empréstimo com ou sem garantias ser o mesmo (Pinheiro *et al.*, 1998), a existência de uma garantia real assegura que haverá ativos para a realização da penhora.

Tudo o mais constante, um empréstimo com garantias pode diminuir as perdas do banco em caso de inadimplência se a sua execução for rápida e pouco custosa. Se isso for verdade, a eficiência do judiciário deve influenciar a recuperação da garantia, mas não deve afetar a taxa de juros (exceto indiretamente através da garantia). Um judiciário pouco eficiente faz com que as garantias sejam muito custosas ou lentas para serem recuperadas e isso se traduz em alta carga de trabalho e/ou alta taxa de congestionamento. Por outro lado, um judiciário mais eficiente contribui para uma maior agilidade na recuperação das garantias. Isso é possível com um maior número de magistrados. Utilizamos como instrumento para a garantia algumas

³⁰ Esta variável foi calculada para o ano de 2000, o mais próximo da data-base dos dados do SCR.

medidas de eficiência do judiciário (*ctrabalho2*, *tcongest2*, *magist* , descritas na **Tabela 2.1** e cujos valores estão no **APÊNDICE 2.A**, p.104). Essas variáveis estão disponíveis para cada estado e, na base do SCR, têm a desvantagem de apresentar pequena variabilidade.

Outra hipótese acerca do instrumento é de que a localização geográfica deve estar correlacionada com a garantia. Assim como o valor dos bens imóveis e móveis pode variar de acordo com a região, a própria eficiência do judiciário dentro do mesmo estado deve ser distinta, se comparar a capital com o interior. Por outro lado, consideramos que a taxa de juros dos empréstimos não deve ser influenciada diretamente pela localização. Dado que utilizamos algumas características da transação como variáveis de controle, as taxas de juros entre contratos semelhantes não devem variar somente porque foram celebrados em regiões geográficas diferentes. Dessa forma, adotamos a hipótese de que a localização da agência que concedeu o empréstimo está correlacionada com a garantia, afetando indiretamente a taxa de juros através do seu efeito sobre a garantia. Incluímos, portanto, como instrumento 54 *dummies* de localização geográfica.

Nas estimações da equação da taxa de juros por *treatment effect* com endogeneidade e *switching regression model*, os instrumentos aqui apresentados são utilizados como restrições de exclusão, ou seja, são as covariadas exógenas da equação binária. A **Tabela 2.1** traz a descrição dos instrumentos utilizados:

Tabela 2.1 – Descrição dos instrumentos

Instrumento	Descrição
<i>hdep</i>	Índice <i>Herfindahl</i> de depósitos calculado para o município em que a agência que efetuou a transação está localizada. É calculado da seguinte forma: $hdep = \sum_{j=1}^n s_j^2$, onde s_j^2 é o quadrado da proporção do volume de depósitos do banco <i>j</i> no município. Fonte: ESTBAN.
<i>cap</i>	Razão do Capital Residencial Total ³¹ R\$(mil) em 2000 dividido pelo Número de Domicílios Particulares Permanentes em 2000, calculado para cada município. O Número de Domicílios Particulares Permanentes é calculado pelo Censo Demográfico do IBGE.

³¹ O Capital Residencial Total, calculado pelo IPEA, é igual ao “valor presente do fluxo perpétuo constantes dos aluguéis mensais descontado à taxa de desconto de 0,75% a.m. O aluguel dos imóveis, inclusive dos próprios, foi simulado por um modelo hedônico cujos argumentos são os atributos dos imóveis, inclusive sua localização, captada pela renda mediana do setor censitário. Em todos os anos utilizou-se as estimativas dos preços hedônicos referentes a 1999 obtendo-se, portanto, um índice de base fixa. Para os anos de 1970, 1980, 1991 e 2000 foram utilizados dados do Censo Demográfico, para os demais, da PNAD”.

Instrumento	Descrição
<i>pot</i>	Razão volume de crédito / PIB calculado para o município em que a agência que efetuou a transação está localizada. Fonte: IBGE.
<i>magist</i>	Magistrados por 100 mil habitantes na justiça estadual. Fonte: CNJ.
<i>ctrabalho2</i>	Carga de trabalho no 2º grau = [(Casos novos no 2º grau + Casos pendentes de julgamento no 2º grau) / Número de magistrados de 2º grau no Estado]. Fonte: CNJ.
<i>tcongest2</i>	Taxa de congestionamento no 2º grau = $1 - [N^{\circ} \text{ de decisões que extinguem o processo no } 2^{\circ} \text{ grau} / (\text{Casos novos no } 2^{\circ} \text{ grau} + \text{Casos pendentes de julgamento no } 2^{\circ} \text{ grau})]$. Fonte: CNJ.
<i>dreg1 a</i> <i>dreg54</i>	Variáveis <i>dummies</i> =1 se a agência bancária que concedeu o empréstimo está localizada na região correspondente e =0 caso contrário. Cada uma das 27 UFs foi dividida em duas partes: capital e resto da UF, totalizando 54 regiões.

2.4 DADOS

O mercado oferta diversas modalidades de empréstimo para pessoas jurídicas – PJ, de acordo com a finalidade de uso, dentre as quais podemos citar: conta garantida, *hot money*, capital de giro, desconto de títulos e duplicatas, *vendor*, *comprar*, financiamento de veículos, *leasing* e financiamento rural.

A modalidade selecionada para este trabalho é o Capital de Giro e, conforme observamos na **Figura 2.1**, é uma das principais operações em termos de volume. No entanto, o principal motivo para a seleção dessa operação se deve ao fato de que a mesma não possui garantias vinculadas diretamente ao empréstimo. Por exemplo, no financiamento de veículos ou no desconto de duplicatas o próprio bem financiado (veículo ou a duplicata, respectivamente) serve como garantia, tornando-se desinteressante para o estudo em questão.

Fonte: Banco Central do Brasil.

O Capital de Giro é uma modalidade de empréstimo que é contratada antes da utilização dos recursos e por isso o cliente pode negociar previamente o valor, os prazos e as taxas. Por parte do banco, a previsibilidade permite um melhor casamento entre a captação e a aplicação de recursos (vide Ortolani (2000) p. 67).

As características de cada uma das modalidades refletem-se diretamente na taxa de juros da operação, conforme observamos na **Figura 2.2**. A Conta Garantida, pelo seu caráter de imprevisibilidade na utilização, possui as maiores taxas. O Capital de Giro possui taxas intermediárias entre a Aquisição de Bens e a Conta Garantida.

Fonte: Banco Central do Brasil.

2.4.1 Base de dados

A principal fonte dos dados utilizados neste trabalho é o SCR – Sistema de Informações de Crédito do Banco Central do Brasil. As complementares são CNJ - Conselho Nacional de Justiça, o IPEA, o IBGE e o ESTBAN e foram utilizadas para compor os instrumentos.

O SCR é uma fonte rica de microdados referentes a crédito e neste trabalho foram utilizadas as operações para pessoas jurídicas na modalidade Capital de Giro. A base traz informações detalhadas e individualizadas de todas as operações de clientes iniciadas em set/2005 cujo valor é superior a R\$ 5.000,00. Como características do contrato temos as seguintes variáveis: a taxa de juros (*juros*), a duração da operação (*duracao*), sua classificação de risco (*risco*) e se o empréstimo possui garantias não fidejussórias (*gar*). Quanto às características do tomador temos o tempo de relacionamento com o banco (*relac*), o número de bancos ativos com os quais mantém relacionamento (*totif*), o número de operações no SCR (*totoper*) e o

porte da empresa (*portecli*). Duas variáveis dummies indicam a capacidade de pagamento do tomador: a primeira se existem valores vencidos no SCR (*dvenc*) e a segunda, se existem valores baixados a prejuízo no SCR (*dprej*). Conseguimos também diferenciar as instituições que concederam crédito (*idif*). A descrição detalhada dos dados está na **Tabela 2.2**:

Tabela 2.2 – Descrição das variáveis

Variável	Descrição
<i>gar</i>	Variável <i>dummy</i> , igual a 1 se foi apresentado algum tipo de garantia não-fidejussória para a concessão do empréstimo e zero caso contrário.
<i>ljuros</i>	Ln (logaritmo natural) da taxa de juros em % a.a..
<i>duracao</i>	Prazo para pagamento do crédito em meses.
<i>relac</i>	Tempo de relacionamento do tomador do empréstimo com a instituição financeira, em anos.
<i>risco</i>	É a classificação de risco da operação e determina a perda mínima esperada para uma operação de crédito. Para operações inferiores a R\$ 50 mil o SCR aceita a classificação de acordo com o atraso. Constitui um dos principais parâmetros para determinar a provisão da operação. Está dividida em nove categorias, que varia de AA a H, sendo a classificação AA considerada a menos arriscada e H a mais arriscada. Na amostra utilizada, selecionamos apenas quatro categorias e atribuímos os seguintes valores: AA=1, A=2, B=3, C=4.
<i>portecli</i>	Porte do tomador: igual 1 se micro, 2 se pequeno, 3 se médio e 4 se grande. Este valor pode ser contínuo, devido à substituição dessa variável pela média da avaliação do porte do cliente realizada por outras IFs quando o valor era <i>missing</i> .
<i>totif</i>	Número de IFs ativos com as quais o tomador está relacionado.
<i>totoper</i>	Total de operações que o tomador possui no SCR/1000.
<i>porteif</i>	Porte da IF, igual 1 se micro, 2 se pequeno, 3 se médio e 4 se grande.
<i>dvenc</i>	Variável <i>dummy</i> igual a 1 se o tomador possui valores vencidos no SCR.
<i>dprej</i>	Variável <i>dummy</i> igual a 1 se o tomador possui valores baixados a prejuízo no SCR.

2.4.2 Análise preliminar dos dados

Com a base de dados completa, para cada cliente, calculamos a média das avaliações do *porte* feita pelas IFs (quando existir) e a substituímos quando *porte* = 0 ou *missing*. Após, selecionamos as operações com taxas pré-fixadas e com recursos próprios da modalidade Capital de Giro. As operações com classificação de risco elevada foram eliminadas (*risco* > 4), pois as mesmas podem contemplar operações renegociadas e, por isso, os termos dos contratos podem ser distintos dos que vigorariam no mercado. Excluímos também as

instituições financeiras que não apresentavam variabilidade na garantia (somente operações com(sem) garantias)³².

Na **Tabela 2.3** e na **Tabela 2.4** apresentamos as estatísticas descritivas das variáveis do SCR, que totalizam 51.617 observações. A maioria dos contratos não possui garantias e 13,88% das operações de Capital de Giro apresentam garantia real. Uma operação média apresenta taxa de juros de 45,63% a.a., com prazo de 5,80 meses. A média da classificação de risco é de 2,38 e a empresa é de porte pequeno a médio. Os tomadores em média se relacionam com aproximadamente 6 instituições financeiras. As variáveis *dvenc* e *dprej* são binárias e iguais a 1 se o tomador possui respectivamente crédito vencido ou baixado a prejuízo no SCR. Como essas variáveis podem assumir valores 1 somente para quem já tomou um empréstimo anterior, elas não servem como indicador da qualidade do crédito para os novos tomadores do sistema. Em média, 18% da amostra possui valores vencidos e 1% valores baixados a prejuízo. A maioria dos empréstimos da amostra foi concedida por um banco de porte grande. Como não existe muita variabilidade no porte da IF (*porteif*), optamos por apresentar os resultados sem incluí-la nas estimativas³³.

Tabela 2.3 – Estatística descritiva das variáveis – Capital de Giro

Variável	Média	Desvio-Padrão	1° quartil	Mediana	3° quartil
<i>juros</i>	45,63	15,90	30,76	42,57	59,73
<i>gar</i>	0,14	0,35	0	0	0
<i>duracao</i>	5,80	4,60	1,70	3,43	12,10
<i>relac</i>	22,30	36,32	1,40	5,48	17,31
<i>risco</i>	2,38	1,05	2	2	3
<i>portecli</i>	2,32	0,69	2	2	3
<i>totif</i>	5,62	3,51	3	5	7
<i>totoper</i>	0,14	0,17	0,04	0,07	0,15
<i>porteif</i>	3,97	0,24	4	4	4
<i>dvenc</i>	0,18	0,39	0	0	0
<i>dprej</i>	0,01	0,11	0	0	0

Nº Observações: 51.617 (set/2005). Fonte: SCR – Bacen.

³² Do total de operações pré-fixadas da amostra, a seguinte ordem de eliminação foi efetuada: 1) valor de empréstimo ≤ 0 (~2%); 2) tempo de relacionamento ≤ 0 (~8%); 3) duração ≤ 0 (~0,3%); 4) duração > 365 (~22%); 5) valor do empréstimo superior a R\$ 1,012 milhões (~0,5%); 6) taxa de juros $> 92,51\%$ a.a. (~0,7%); 7) taxa de juros ≤ 0 (~0,2%); 8) porte do cliente *missing* (~8%); 9) classificação de risco > 4 (~3%); e 9) porte da IF *missing* (~3%); restando 51.617 observações.

³³ Incluí-la ou não não impacta significativamente os coeficientes estimados.

Na **Tabela 2.4**, alguns padrões são evidentes. Em média, uma operação com garantia tem taxa de juros menor que as sem garantia. Nas operações com garantia, os clientes possuem tempo de relacionamento mais curto (média de 5,34 anos), se relacionam com mais IFs (em média, 7), a duração do empréstimo e o risco são em média maiores, bem como os clientes apresentam maior taxa de inadimplência (créditos vencidos/baixados a prejuízo).

Para verificar se existem diferenças estatísticas acerca das variáveis apresentadas na **Tabela 2.4** procedemos ao teste de diferença de médias. A hipótese nula (H_0) a ser testada é se inexistente diferença entre os valores médios das variáveis dos empréstimos com e sem garantia, considerando ambas as amostras independentes. A hipótese alternativa (H_a) é que elas são diferentes:

$$H_0 : \mu_1 = \mu_0 \quad \text{e} \quad H_a : \mu_1 \neq \mu_0$$

Onde μ_1 e μ_0 são as médias populacionais dos empréstimos com e sem garantia, respectivamente. A estatística de teste Z ³⁴ pode ser aproximada por uma distribuição normal e para um nível de significância de 5% bicaudal ($Z_{0,025} = 1,96$). Rejeitamos a hipótese nula para todas as variáveis, exceto para o porte do cliente.

³⁴ A estatística do teste é calculada como: $Z = \frac{(\bar{x}_1 - \bar{x}_0) - (\mu_1 - \mu_0)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_0^2}{n_0}}}$, onde $\bar{x}$ é a média amostral, s^2 é o

desvio-padrão amostral e n é o tamanho da amostra. Os índices 1 e 0 representam os empréstimos com e sem garantia, respectivamente.

Tabela 2.4 – Estatística descritiva das variáveis por garantia – Capital de Giro

Variável	gar=1 N° obs: 7.164 (13,9%)		gar=0 N° obs: 44.453 (86,1%)		Teste de Média $H_0: \mu_1 = \mu_0$
	Média	Desvio-Padrão	Média	Desvio-Padrão	Estat-Z
<i>juros</i>	39,57	10,01	46,61	16,44	-49,70
<i>duracao</i>	8,56	3,97	5,36	4,54	62,00
<i>relac</i>	5,34	14,98	25,04	37,97	-78,02
<i>risco</i>	2,53	0,87	2,36	1,08	14,80
<i>portecli</i>	2,31	0,85	2,32	0,67	-0,95
<i>totif</i>	6,52	4,23	5,48	3,35	19,83
<i>totoper</i>	0,09	0,10	0,14	0,18	-34,30
<i>dvenc</i>	0,23	0,42	0,18	0,38	9,47
<i>dprej</i>	0,02	0,14	0,01	0,10	5,81

Fonte dos dados: SCR – Bacen (set/2005).

2.5 RESULTADOS EMPÍRICOS

Nesta seção, apresentamos os resultados. Inicialmente, mostramos as estimativas da equação da garantia, seguida pela equação da taxa de juros. Sempre que possível, apresentamos os resultados obtidos por outros autores que utilizaram técnicas similares. Para verificar a robustez, estimamos novamente as regressões considerando apenas uma observação por tomador.

2.5.1 Equação da garantia

Na **Tabela 2.5** apresentamos os resultados da equação da garantia. Respectivamente na coluna (1) e nas (2)-(3), reportamos a estimação sem considerar e considerando a endogeneidade da taxa de juros. Nos três casos o sinal da variável *ljuros* é negativo e significativo, evidenciando que os contratos de crédito com menor taxa de juros são aqueles que apresentam garantias. Quando endógena, a taxa de juros foi instrumentalizada pelo Índice de *Herfindahl* de depósitos, pela razão crédito/PIB e por *dummies* de IFs. O valor (absoluto) deste coeficiente é maior para os modelos que consideram a endogeneidade.

Nas três especificações, os valores dos coeficientes das variáveis exógenas são próximos, diferindo apenas em alguns casos na magnitude e na significância. *Risco* é positivo e significativo nas estimações que tratam a taxa de juros como endógena. Os coeficientes de *portecli* e *dvenc* impactam muito mais (em módulo) a probabilidade de o empréstimo possuir garantias quando se considera a endogeneidade.

Duracao e *risco* são as *proxies* do risco num contrato de crédito. Ambas variáveis apresentam sinais positivos, ou seja, quanto mais arriscada é a operação, maior a probabilidade de o empréstimo apresentar garantias.

A variável associada ao tempo de relacionamento (*relac*) apresenta coeficiente negativo e significativo, ou seja, existe evidência de que um tempo de relacionamento mais longo ajuda a diminuir os efeitos da assimetria informacional. Como ela deve ser determinada conjuntamente com o número de operações que o tomador possui (*totoper*) e o total de IFs

com as quais mantém relacionamento (*totif*), essas variáveis servem de controle para *relac*. *Totif* apresenta coeficiente negativo e significativo apenas na especificação (1), quando não se considera a endogeneidade, e nas demais, este coeficiente não é significativo. O coeficiente de *totoper* é sempre negativo e significativo.

Tabela 2.5 – Resultados das estimativas do modelo *probit*

Variável	(1)			(2)			(3)		
	Coef.	Desv. Padr.		Coef.	Desv. Padr.		Coef.	Desv. Padr.	
<i>ljuros</i>	-1,6749	0,0335	*	-4,4096	0,0465	*	-4,4344	0,0491	*
<i>duracao</i>	0,0717	0,0020	*	0,0690	0,0023	*	0,0670	0,0023	*
<i>relac</i>	-0,0147	0,0008	*	-0,0115	0,0004	*	-0,0123	0,0004	*
<i>risco</i>	0,0071	0,0076		0,0438	0,0086	*	0,0406	0,0088	*
<i>portecli</i>	-0,0558	0,0151	*	-0,6180	0,0160	*	-0,5742	0,0169	*
<i>totif</i>	0,0717	0,0045	*	-0,0005	0,0039		0,0029	0,0044	
<i>totoper</i>	-2,6177	0,2798	*	-1,8473	0,1970	*	-2,0441	0,2278	*
<i>dvenc</i>	0,0870	0,0200	*	0,3256	0,0170	*	0,3196	0,0174	*
<i>dprej</i>	0,2081	0,0630	*	0,2381	0,0627	*	0,1972	0,0614	*
<i>_cons</i>	4,0660	0,2976	*	16,9516	0,2129	*	16,9407	0,2978	*
<i>/athrho</i>	-	-		1,2809	0,0423	*	1,2348	0,0427	*
<i>Dummies:</i>	<i>regiões</i>			-			<i>regiões</i>		
<i>Instrumentos:</i>	-			<i>hdep, pot, dummies IFs</i>			<i>hdep, pot, dummies IFs</i>		

A variável dependente é *gar*. Nº de observações: 51.564 (1) e (3) / 51.617 (2).

Desvio-padrão robusto à heterocedasticidade. Coeficientes das *dummies* de regiões não reportados.

As colunas (2) e (3) consideram endogeneidade da taxa de juros. Teste de *Wald* para exogeneidade (*/athrho*=0).

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Se a IF pode observar o risco do tomador, então aos de baixa qualidade devem ser ofertados taxa de juros mais elevada (e igualmente, solicitado garantia). *Portecli*, *dvenc* e *dprej* são as variáveis que o banco pode utilizar para reduzir a assimetria informacional. Quanto menor o porte do cliente, mais as IFs têm dificuldade para obter informações (opacidade informacional), resultando em empréstimos mais arriscados. Uma forma disponível ao banco para verificar o tipo do tomador é consultando no SCR se o mesmo possui empréstimos baixados à prejuízo ou vencidos³⁵ anteriores à data da concessão do crédito. Os sinais desses coeficientes são significativos e estão de acordo com o esperado: *portecli* tem sinal negativo e *dvenc* e *dprej*, positivos.

Nas duas últimas colunas da **Tabela 2.5**, o coeficiente */athrho* é uma transformação de *rho* que é utilizado no processo de estimação. *Rho* mede a correlação entre os erros da equação do

³⁵ A IF precisa de autorização expressa do tomador para acessar os seus dados no SCR.

probit e a equação da forma reduzida da variável endógena. A hipótese nula é que *ljuros* é exógena. Para as estimativas das colunas (2) e (3), este coeficiente é significativo e, portanto, rejeitamos a hipótese que a taxa de juros seja exógena no modelo.

Os coeficientes estimados na **Tabela 2.5** são difíceis de interpretar e foram utilizados para analisar os sinais e as significâncias. A investigação dos determinantes da probabilidade do empréstimo apresentar garantias é completada pelos efeitos marginais da variação nas explicativas. Para as estimações apresentadas, reportamos estes efeitos calculados para as médias das variáveis independentes (vide **Tabela 2.6**):

Tabela 2.6 – Efeitos marginais do modelo *probit*

Variável	(1)		(2)		(3)	
	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.
<i>ljuros</i>	-0,2191	0,0062 *	-1,0565	0,0309 *	-1,0163	0,0327 *
<i>duracao</i>	0,0094	0,0004 *	0,0165	0,0004 *	0,0153	0,0004 *
<i>relac</i>	-0,0019	0,0001 *	-0,0028	0,0001 *	-0,0028	0,0001 *
<i>risco</i>	0,0009	0,0010	0,0105	0,0021 *	0,0093	0,0021 *
<i>portecli</i>	-0,0073	0,0020 *	-0,1481	0,0062 *	-0,1316	0,0063 *
<i>totif</i>	0,0094	0,0005 *	-0,0001	0,0010	0,0007	0,0010
<i>totoper</i>	-0,3424	0,0326 *	-0,4426	0,0431 *	-0,4685	0,0474 *
<i>dvenc</i> [@]	0,0119	0,0029 *	0,0860	0,0055 *	0,0810	0,0055 *
<i>dprej</i> [@]	0,0316	0,0110 *	0,0637	0,0185 *	0,0498	0,0170 *
<i>Dummies:</i>	<i>regiões</i>		-		<i>regiões</i>	
<i>Instrumentos:</i>	-		<i>hdep, pot, dummies IFs</i>		<i>hdep, pot, dummies IFs</i>	

A variável dependente é *gar*. N° de observações: 51.564 (1) e (3) / 51.617 (2). Efeitos marginais das *dummies* de regiões não reportados. As colunas (2) e (3) consideram endogeneidade da taxa de juros.

[@]: o efeito marginal é calculado para uma variação discreta da *dummy* de 0 para 1.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Na coluna (1), o coeficiente *ljur* é interpretado da seguinte maneira: para um empréstimo avaliado no valor médio de suas características, se a taxa de juros aumentar em 1%, a probabilidade de o empréstimo ter garantia cai em 0,22. Esse valor torna-se muito maior, em módulo, quando se considera a endogeneidade da taxa de juros, passando para -1,02³⁶.

Procedendo a análise apenas dos efeitos marginais da coluna (3) (nas demais é análoga), verifica-se que o aumento em um ano no tempo de relacionamento diminui a chance de um contrato de crédito médio possuir garantias em 0,28%. Alargar o prazo de pagamento em um mês aumenta a probabilidade de garantia em 1,53% e aumentar a classificação de risco da

³⁶ A inclinação de uma função pode ser maior do que 1, mesmo que os seus valores estejam entre 0 e 1. A linha da tangente é uma boa aproximação da curva próximo ao ponto de tangência, mas se a curva muda rapidamente, essa aproximação não é muito boa longe deste ponto (vide: http://www.stata.com/support/faqs/stat/mfx_size.html, acesso em 18/08/2008).

operação impacta a probabilidade de garantia em 0,93%. $D_{venc}=1$ e $d_{prej}=1$ tornam o crédito mais provável de possuir garantia em 8 e 5%, respectivamente.

Os determinantes do uso da garantia nos contratos de pequenas e médias empresas japonesas é o objeto de investigação de Ono e Uesugi (2008). *Coll* e *guar* são duas variáveis binárias iguais a 1 se o empréstimo possui garantia real e pessoal, respectivamente, e 0 caso contrário. *Rate* é a taxa de juros e *score* é a classificação de crédito da empresa medida pelo TSR – *Tokyo Shoko Research*. A estimação de *coll* como variável dependente, *guar* como endógena, *rate*³⁷ e *score* como independentes e outros controles apresentaram os seguintes resultados: 1) a variável *rate* é positiva e significativa; 2) *score* é não significativo indicando que o risco da firma não tem efeitos sobre a garantia; e 3) um forte relacionamento entre o tomador e o credor contribui para a utilização da garantia.

Jiménez *et al.* (2004) estimam os determinantes das garantias nos contratos de crédito (através do teste de seis hipóteses) usando os dados da *Central de Información de Riesgos* do Banco Central Espanhol. Define-se *collateral* igual a 1 se o empréstimo possui garantia e 0 caso contrário. Estima-se o modelo por *logit* tendo *collateral* como variável dependente, variáveis exógenas que capturam o risco de crédito *ex ante* (*dummy* igual a 1 para os que possuíam valores em atraso) e *ex post* (*dummy* igual a 1 para os que deram default após tomar o crédito) do tomador e a intensidade de relacionamento, entre outras, e diversos controles. Os autores concluem que o uso da garantia aumenta com o risco de crédito devido ao fato de que grande proporção dos empréstimos é feita sob o risco observado e não sob informação privada. A probabilidade de o empréstimo possuir garantia diminui com o relacionamento (tempo, escopo e número de bancos). Em outro trabalho similar usando também dados espanhóis, Jiménez e Saurina (2004) investigam qual o impacto da garantia na probabilidade dos empréstimos bancários sofrerem default. Os autores encontram evidências de que o uso da garantia aumenta as chances de inadimplência.

Os resultados apresentado na **Tabela 2.5** e **Tabela 2.6** diferem dos apresentados por esses autores, exceto para o tempo de relacionamento (Jiménez *et al.*, 2004). No nosso caso, o fato de uma maior taxa de juros diminuir a probabilidade de o empréstimo apresentar garantias é um indício de que neste mercado predomina o efeito da assimetria informacional. Além disso,

³⁷ O teste indicou a não rejeição da hipótese nula de que a variável *rate* é exógena.

o tempo de relacionamento ajuda a diminuir os problemas decorrentes dessa assimetria. A seguir, verificamos se este resultado se mantém quando a garantia passa a ser variável explicativa e a taxa de juros, a variável explicada.

2.5.2 Equação da taxa de juros

Nesta subseção, apresentamos os resultados da estimação por OLS e 2SLS da equação da taxa de juros, cujos resultados estão apresentados nas colunas (1)-(2) e (3)-(6) da **Tabela 2.7**, respectivamente. As variáveis utilizadas como instrumento são: o capital residencial médio (*cap*), um indicador de eficiência do judiciário (*ctrabalho2*) e *dummies* de regiões. Nos resultados das colunas (2), (5) e (6) controlamos para os efeitos fixos das IFs através da inclusão das *dummies*. Em todos os casos, o coeficiente da variável *gar* é negativo e significativo, diferindo apenas na sua magnitude. Em particular, nas estimativas por OLS, o valor deste coeficiente é sempre viesado para baixo, se comparado com o 2SLS. Nas colunas (2) e (6), os coeficientes -0,17 e -0,25 são interpretados da seguinte forma: se um empréstimo sem garantias passa a apresentar garantias, a taxa de juros diminuiria em 16% e 22%³⁸, respectivamente.

Os coeficientes de *duracao* e *risco* ora apresentam sinal positivo, ora sinal negativo, dependendo da especificação. O coeficiente da variável *relac* é negativo e significativo (exceto na coluna (5)), sendo um indício em favor da tese de que a assimetria informacional é reduzida com o tempo de relacionamento. Seu efeito está controlado pelas variáveis *totif* e *totoper*. Conforme esperado, o porte do tomador influi negativamente na taxa de juros. A variável *dummy dvenc* apresenta sinal positivo e significativo, *dprej*, por sua vez não é significativo.

³⁸ $e^{-0,17} - 1 = -0,16$ e $e^{-0,25} - 1 = -0,22$.

Tabela 2.7 – Resultados das estimativas da equação da taxa de juros – OLS e 2SLS

Variável	OLS (1)	OLS (2)	2SLS (3)	2SLS (4)	2SLS (5)	2SLS (6)
<i>gar</i>	-0,1984 * (0,0031)	-0,1669 * (0,0039)	-1,3162 * (0,1041)	-0,2354 * (0,0186)	-0,8682 * (0,1820)	-0,2540 * (0,0420)
<i>duracao</i>	0,0102 * (0,0003)	-0,0063 * (0,0003)	0,0274 * (0,0017)	0,0108 * (0,0004)	0,0008 (0,0019)	-0,0054 * (0,0005)
<i>relac</i>	-0,0016 * (0,0000)	-0,0001 * (0,0000)	-0,0030 * (0,0001)	-0,0017 * (0,0001)	0,0001 (0,0001)	-0,0001 * (0,0000)
<i>risco</i>	-0,0319 * (0,0011)	0,0141 * (0,0012)	-0,0272 * (0,0019)	-0,0317 * (0,0011)	0,0203 * (0,0024)	0,0149 * (0,0013)
<i>portecli</i>	-0,1650 * (0,0023)	-0,1266 * (0,0022)	-0,1276 * (0,0057)	-0,1638 * (0,0024)	-0,0777 * (0,0131)	-0,1205 * (0,0036)
<i>totif</i>	-0,0110 * (0,0004)	-0,0106 * (0,0003)	0,0056 * (0,0017)	-0,0104 * (0,0005)	-0,0055 * (0,0014)	-0,0100 * (0,0005)
<i>totoper</i>	-0,2912 * (0,0104)	-0,0003 (0,0078)	-0,4110 * (0,0203)	-0,2952 * (0,0106)	-0,0842 * (0,0247)	-0,0107 (0,0090)
<i>dvenc</i>	0,0759 * (0,0032)	0,0462 * (0,0028)	0,0642 * (0,0058)	0,0756 * (0,0032)	0,0413 * (0,0039)	0,0456 * (0,0028)
<i>dprej</i>	0,0146 (0,0122)	-0,0056 (0,0111)	0,0927 * (0,0259)	0,0172 (0,0124)	0,0000 (0,0145)	-0,0049 (0,0112)
<i>_cons</i>	4,3113 * (0,0061)	4,8234 * (0,0176)	4,2227 * (0,0136)	4,3084 * (0,0062)	5,2949 * (0,1281)	4,8820 * (0,0341)
<i>Dummies:</i>	-	IFs	-	-	IFs	IFs
<i>Instrumentos:</i>	-	-	<i>cap,</i> <i>ctrabalho2</i>	<i>dummies</i> <i>regiões, cap,</i> <i>ctrabalho2</i>	<i>cap,</i> <i>ctrabalho2</i>	<i>dummies</i> <i>regiões, cap,</i> <i>ctrabalho2</i>
<i>Estatística-F</i> <i>do 1º</i> <i>Estágio:</i>	-	-	699,49	133,95	2.628,50	1.340,88

A variável dependente é *ljuros*. N° de observações: 51.617.

Desvio-padrão robusto à heterocedasticidade entre parênteses. Coeficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

A troca do instrumento *ctrabalho2* por *magist* ou *tcongest2* mantém os resultados na maioria dos casos.

A estimativa por OLS do coeficiente da garantia, na segunda coluna da **Tabela 2.7** (-0,17) deve ser inconsistente se *gar* é uma variável endógena. Estimando por 2SLS, este coeficiente aumenta, em módulo, para -0,25 (coluna (6)). No entanto, nota-se que a estimativa por 2SLS é menos eficiente que por OLS devido à magnitude do desvio-padrão robusto à heterocedasticidade. No primeiro caso esse valor é de 0,04 (coluna (6)), enquanto no segundo é 0,004 (coluna (2)), aproximadamente dez vezes maior. Para as variáveis exógenas, essa diferença é bem menor³⁹. O R^2 -parcial (*Shea partial R²*) (coluna (6)) é bastante pequeno

³⁹ As variáveis exógenas apresentam um instrumento bastante forte, pois a própria variável serve como instrumento para ela mesma.

(=0,007) levando à perda na precisão da estimação⁴⁰. Isto explica porque o erro-padrão do coeficiente estimado por 2SLS pode ser muito maior que o OLS, sendo um indício de que os instrumentos utilizados são fracos (Cameron e Trivedi, 2005, capítulo 4). Resultado análogo é obtido para as colunas (3)-(5).

Outra forma de checar se os instrumentos são fracos quando existe um único regressor endógeno é computando a estatística-F, que testa a hipótese de que todos os coeficientes são iguais a zero no primeiro estágio da regressão de 2SLS. Esta estatística fornece uma medida do conteúdo informacional dos instrumentos e, quanto maior, maior é o seu valor esperado. Uma “regra de bolso” utilizada é que se a estatística-F do primeiro estágio exceder 10, não há necessidade de se preocupar com os instrumentos fracos (Stock e Watson, 2003, p.350). Essa condição é atendida para as quatro estimações apresentadas.

Adicionalmente, realizamos também o teste de sobreidentificação, cuja hipótese nula conjunta é de que todos os instrumentos são válidos. Todos os modelos apresentados rejeitam a hipótese nula, confirmando o resultado anterior. Contudo, a estatística deste teste é proporcional ao número de observações e a sua rejeição pode ser devido à grande quantidade de observações da nossa amostra.

A partir dos resultados desses testes não temos evidências fortes de que os instrumentos utilizados apresentam as características desejáveis. Nossa motivação para a sua utilização é mais teórica do que empírica. De certa forma, estamos mais interessados no sinal da garantia do que na sua magnitude e ambas as metodologias utilizadas apontam para a mesma direção.

Os resultados obtidos por outros autores são apresentados a seguir. Conforme já mencionado, Berger e Udell (1990 e 1995) e Degryse e Van Cayseele (2000) consideram que o processo de determinação da garantia e da taxa de juros é seqüencial, com a garantia precedendo a taxa de juros.

⁴⁰ A variância assintótica da estimação 2SLS é: $se[\hat{\beta}_{1,2SLS}] = \frac{se[\hat{\beta}_{1,OLS}]}{R_p}$, onde $\hat{\beta}_1$ é o coeficiente do regressor

endógeno, R_p^2 é o R^2 -parcial “limpo” dos instrumentos exógenos. Com instrumentos fracos, o erro-padrão do 2SLS aumenta devido ao R^2 -parcial ser baixo.

Berger e Udell (1990) usam dados do *Federal Reserve's Survey of Terms of Banking Lending* e estimam uma regressão usando prêmio de risco do empréstimo como variável dependente, medidas de garantia como exógena e diversos controles. O resultado obtido sugere que o paradigma do risco observado ocorre mais freqüentemente que o paradigma da informação privada.

O papel do relacionamento nos empréstimos para firma de pequeno porte é o tema de estudo de Berger e Udell (1995), usando os dados do *National Survey of Small Business Finance* (NSSBF). Duas equações são estimadas: 1) uma com o prêmio de risco como variável dependente, medidas de garantia e relacionamento como exógenas e diversos controles; 2) outra com a variável binária *collat* (=1 se o empréstimo possui garantia e 0 caso contrário) como dependente, relacionamento como exógena e diversos controles. A garantia não é estatisticamente significativa na equação da taxa de juros. No entanto, as evidências apontam que, para as firmas de pequeno porte, um maior tempo de relacionamento implica menor taxa de juros e menor probabilidade de garantia no empréstimo.

Degryse e Van Cayseele (2000) estudam o impacto do relacionamento (duração e escopo) entre o cliente e a instituição financeira nos termos do contrato de empréstimo. A base de dados compreende empréstimos de empresas pequenas belgas. Duas equações são especificadas e estimadas, sendo uma para a taxa de juros (OLS) e outra para a garantia (*logit*). Os autores encontram indícios de que o tempo de relacionamento influi positivamente e o escopo (quantidade de produtos e serviços comprados do banco) negativamente na taxa de juros, sendo o sinal desta mais importante. O sinal negativo da garantia na equação da taxa de juros é compatível com o paradigma da informação privada. A garantia, por sua vez, é influenciada negativamente pelo tempo e positivamente pelo escopo.

Nosso resultado com relação ao papel da garantia na determinação da taxa de juros é compatível com o obtido por Degryse e Van Cayseele (2000). Quanto ao tempo de relacionamento, está em consonância com os apresentados por Berger e Udell (1995) e Degryse e Van Cayseele (2000), apesar de nosso trabalho não estar restrito a pequenas empresas.

2.5.3 *Treatment effects*

Outra técnica utilizada para estimar a equação da taxa de juros é o efeito do tratamento (*treatment effect*). Neste caso, queremos saber qual é a variação média na taxa de juros se um contrato sem garantias passa a apresentar garantias⁴¹. Como não podemos observar o mesmo empréstimo com e sem garantias, nos atemos ao efeito médio do tratamento.

O *matching estimator* estima o efeito médio do tratamento sem considerar a endogeneidade da garantia. Na ausência de um fato contrário (*counterfactual*), este estimador procura na amostra indivíduos similares que não foram submetidos ao tratamento. Se a decisão de apresentar garantias é aleatória, podemos usar a média das taxas de juros de contratos similares que não apresentam garantias para estimar a taxa de juros dos não-tratados. Para a estimação do ATE (efeito médio do tratamento), ATT (efeito médio do tratamento para os tratados) e ATC (efeito médio do tratamento para o grupo de controle) utilizamos o *bias-corrected matching estimator*, através do comando *nnmatch* programado por Abadie *et al.* (2004) para o *software* econométrico Stata. Para cada observação, especificamos o número de pares igual a 4. Para tornar o processo de estimação mais rápido, selecionamos aleatoriamente 25% da nossa amostra inicial, totalizando 12.904 observações. Os resultados estão na **Tabela 2.8**. A variável que indica o tratamento (*gar*) é igual a 1 se o empréstimo possui garantia fidejussória e 0 caso contrário. As demais variáveis explicativas são *duracao*, *relac*, *risco*, *portecli*, *totif*, *totoper*, *porteif*, *dvenc* e *dprej*. Na linha (2), além dessas variáveis incluímos também as *dummies* para IFs.

Em todas as estimativas, os coeficientes são negativos e significativos, diferindo apenas na magnitude do valor. Os coeficientes do ATT são os menores e do ATC, os maiores. Na estimação (2), o resultado pode ser interpretado da seguinte maneira: para o ATE (-0,14), se sortearmos aleatoriamente uma operação e passarmos da situação sem garantia para a situação com garantia, haveria redução de -0,14 no logaritmo da taxa de juros. Para as operações com garantia (ATT), o efeito seria a redução no *log* da taxa de juros de -0,07. Analogamente, o efeito médio do tratamento para o grupo de controle (ATC) seria o de diminuir em -0,15 o *log* da taxa de juros.

⁴¹ Denominamos aqui como tratamento uma operação com garantia e não-tratamento as operações sem garantia.

Tabela 2.8 – Resultados das estimativas do *matching estimator*

Estimação		ATE	ATT	ATC
Coeficiente	(1)	-0,1526 *	-0,1359 *	-0,1552 *
Desvio-padrão		0,0119	(0,0076)	(0,0134)
Coeficiente	(2)	-0,1397 *	-0,0718 *	-0,1506 *
Desvio-padrão		(0,0285)	(0,0143)	(0,0328)

A variável dependente é *ljuros*. N° de observações: 25% da amostra=12.904.

Estimação usando o comando *nnmatch* (Abadie *et al.*, 2004).

(2) inclui *dummies* de IFs. Desvio-padrão com correção de viés.

* variável associada significativa a 1%.

Por outro lado, se considerarmos que a presença de garantia é endógena, estimamos o efeito médio do tratamento com variáveis instrumentais. Idealmente, para ajudar na identificação e tornar a estimação mais robusta, seria necessária a existência de variáveis na equação de tratamento que não estivessem presentes na equação da taxa de juros. Na coluna (1) da **Tabela 2.9**, apresentamos o resultado sem considerar nenhuma restrição de exclusão, utilizando apenas a não-linearidade do *probit* para a identificação do modelo. Na coluna (2), utilizamos as variáveis *cap* e *ctrabalho2* e na (3), as *dummies* de regiões e *ctrabalho2*⁴². Na **Tabela A - 2.2** do **APÊNDICE 2.B**, p.105 encontra-se os resultados da equação de seleção.

O tratamento (existência de garantia no contrato de crédito) é estimado como um *probit*. *Rho* e *sigma* representam a correlação entre o erro da equação de tratamento e da taxa de juros e o erro-padrão da equação da taxa de juros, respectivamente. Nas três especificações apresentadas, o coeficiente *rho* é significativo indicando a presença do viés de seleção amostral.

⁴² Em (3), *cap* não é significativo e por isso foi excluído.

Tabela 2.9 – Resultados das estimativas do modelo de efeitos do tratamento

Eq. ln(juros)	(1)		(2)		(3)	
	Coeficiente	Desv. Padr.	Coeficiente	Desv. Padr.	Coeficiente	Desv. Padr.
<i>gar</i>	-0,2234	0,0088 *	-0,5623	0,0263 *	-0,2274	0,0086 *
<i>duracao</i>	-0,0057	0,0003 *	0,0158	0,0005 *	-0,0057	0,0003 *
<i>relac</i>	-0,0001	0,0000 *	-0,0021	0,0001 *	-0,0001	0,0000 *
<i>risco</i>	0,0146	0,0012 *	-0,0303	0,0012 *	0,0146	0,0012 *
<i>portecli</i>	-0,1227	0,0023 *	-0,1528	0,0029 *	-0,1224	0,0023 *
<i>totif</i>	-0,0102	0,0003 *	-0,0056	0,0006 *	-0,0102	0,0003 *
<i>totoper</i>	-0,0071	0,0077	-0,3302	0,0126 *	-0,0075	0,0077
<i>dvenc</i>	0,0458	0,0028 *	0,0721	0,0036 *	0,0457	0,0028 *
<i>dprej</i>	-0,0052	0,0111	0,0400	0,0149 *	-0,0051	0,0112
<i>cons</i>	4,8613	0,0193 *	4,2825	0,0072 *	4,8641	0,0194 *
<i>rho</i>	0,1686	0,0229 *	0,7287	0,0384 *	0,1838	0,0226 *
<i>sigma</i>	0,1991	0,0009 *	0,2771	0,0038 *	0,1992	0,0009 *
<i>Dummies</i>	IFs		-		IFs	
<i>Restrição exclusão:</i>	-		<i>cap, ctrabalho2</i>		<i>dummies regiões, ctrabalho2</i>	

Nº de observações: 51.617. Desvio-padrão robusto à heterocedasticidade.

Coeficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

A troca do instrumento *ctrabalho2* por *magist* ou *tcongest2* não alteram significativamente os resultados.

Na equação do *ln* da taxa de juros, o sinal do coeficiente da variável *gar* é negativo e significativo. Nas especificações (1) e (3) que incluem as *dummies* de IFs, este coeficiente é igual a -0,22 e, na coluna (2), sem essas *dummies*, o seu valor é -0,56. Os coeficientes estimados em (1) e (3) são bastante similares. O valor -0,22 é interpretado da seguinte forma: quando a variável *gar* passa de 0 a 1, o *ln* da taxa de juros reduz-se em 0,22, e a evidência de que, nesse mercado, está presente a assimetria de informação se mantém. Comparando este resultado com o ATE da **Tabela 2.8**, verifica-se que, ali, os valores dos coeficientes são menores em módulo que neste caso (-0,22 contra -0,14 e -0,56 contra -0,15, com e sem as *dummies* de IFs, respectivamente).

A maturidade do empréstimo (*duracao*) e o *risco* apresentam sinais contrários, dependendo da especificação. O coeficiente do tempo de relacionamento novamente é negativo e significativo, sendo um indício de que uma relação mais longa contribui para que o banco acumule informações acerca do tomador, reduzindo a assimetria. *portecli* apresenta sinal negativo e *dvenc*, positivo. *Dprej* é positivo e significativo na coluna (2). De modo geral, os resultados obtidos anteriormente se mantêm.

Utilizando igualmente a base do SCR, no entanto para a *cross-section* de dez/2003, Rodrigues *et al.* (2004) estimam os efeitos médios do tratamento (empréstimo apresentar garantias) sob as hipóteses do *switching regression model*. Para a estimação da equação de seleção, os autores utilizam como função de distribuição o *logit* (capital de giro) e o *logit* assimétrico (conta garantida). Os resultados da estimativa para a modalidade capital de giro foram: -0,73 (ATE), -2,58 (ATT) e -0,22 (ATC). Nosso resultado é compatível com o apresentado por esses autores.

Até o momento temos evidências de que o efeito predominante no mercado de crédito brasileiro é o de substituição entre taxa de juros e garantia. Segundo a teoria apresentada por Bester (1985), neste mercado estaria presente a assimetria informacional. Os bancos ofertariam dois tipos de contratos: 1) com garantia e taxa de juros baixa; e 2) sem garantia e taxa de juros alta. Os clientes com menor probabilidade de inadimplência selecionam o primeiro contrato e os clientes com maior probabilidade de inadimplência, o segundo.

O tempo de relacionamento (negativo e significativo) também indica que o efeito predominante é o da redução da assimetria informacional e não de *hold-up*. Os coeficientes das variáveis *portecli* e *dvenc* indicam que caso seja possível diminuir os efeitos da assimetria (*ex-ante*), é possível reduzir a taxa de juros (ou a probabilidade de o empréstimo apresentar garantia). Esses resultados empíricos estão controlados pelo risco da operação.

Evidentemente, não podemos tomar esses resultados como um teste direto de existência de assimetria informacional (seleção adversa e/ou risco moral). Não acompanhamos o pagamento dos contratos e por isso não temos informações acerca das inadimplências e, portanto, não sabemos se as qualidades *ex-ante* observadas do tomador estão de acordo com as *ex-post*.

No entanto, se for verdade que uma relação negativa entre taxa de juros e garantia é um indício de existência de equilíbrio separador (tomadores de baixo risco escolhem contratos com garantia e baixa taxa de juros), então os clientes que selecionam os contratos podem ser diferentes com relação ao risco não-observado (*ex-ante*). Neste caso, é possível que o efeito da garantia não seja somente no intercepto da taxa de juros, mas que os coeficientes das regressões sejam diferentes para os contratos com e sem garantia. Por isso, propomos o *switching regression model*.

2.5.4 *Switching regression model*

Se acreditarmos que, além do intercepto, outros coeficientes também são distintos entre os grupos, então pode-se utilizar a técnica do *switching regression model*⁴³. Este modelo é indicado quando se admite que os tomadores selecionam os diferentes contratos ofertados baseados em características não observáveis que estão correlacionadas com o tipo do contrato (com e sem garantia) que escolhem.

Tecnicamente, a identificação pode ser feita através da não-linearidade e a equação de seleção irá conter as mesmas variáveis que a equação da taxa de juros. Para tornar a estimação mais robusta seriam necessárias variáveis extras na equação de seleção que não estivessem presentes na equação contínua. Na **Tabela 2.10**, apresentamos os resultados das estimativas da equação da taxa de juros (e na **Tabela A - 2.3** do **APÊNDICE 2.B**, p.105, os resultados da equação binária). Na coluna (1), apresentamos o resultado sem considerar restrição de exclusão e, na coluna (2), utilizamos as variáveis *cap*, *ctrabalho2* as *dummies* de região para auxiliar na identificação do modelo.

Na equação da taxa de juros, o coeficiente do tempo de relacionamento é negativo e do risco é positivo, mas significativos apenas para os empréstimos sem garantia (*gar* = 0). Quanto maior o porte do cliente, menor é a taxa de juros. Tomadores que possuem valores vencidos no SCR têm taxa de juros maiores nos contratos. Em geral, os resultados estão consistentes com os já apresentados, exceto para a variável *duracao*, que apresenta coeficiente negativo e significativo em ambas as especificações.

O coeficiente da covariância entre o erro da equação de seleção e a equação da taxa de juros, *sigma*, é significativo indicando que a mudança de regime é endógena. O resultado de maior interesse é o coeficiente de correlação entre u_{ji} e u_i , $j=1,2$, denominado *rho*. Na especificação (2), este coeficiente é positivo e significativo. O coeficiente positivo para os empréstimos com garantia implica seleção negativa (vide **Equação 2.11**, p.71). Isto significa que os empréstimos com garantias possuem características não-observáveis que reduzem a

⁴³ O método de estimação apresentado nesta subseção está baseado em um pressuposto forte, que é a normalidade conjunta dos termos de erro da equação binária e contínua.

taxa de juros se comparados com um contrato aleatório da amostra. De forma análoga, o efeito da seleção é positivo para os contratos com garantia, o que implica que esses contratos possuem características não-observáveis que aumentam a taxa de juros com relação a um aleatório da amostra. Na especificação (1), a seleção é negativa para $gar = 1$ e não-significativa para $gar = 0$. Nos dois casos, o teste de Wald rejeita a hipótese de que os erros das equações sejam independentes, em favor da distribuição normal conjunta. Adicionalmente, realizamos o teste de igualdade de coeficientes da equação da taxa de juros para $gar = 1$ e $gar = 0$. Em ambos os casos, o teste de Wald ($Prob > \chi^2 = 0,00$) rejeita a igualdade dos coeficientes entre os contratos com e sem garantia.

Tabela 2.10 – Resultado do *endogenous switching regression model*

<i>ljuros</i>	(1)				(2)			
	<i>gar=1</i>		<i>gar=0</i>		<i>gar=1</i>		<i>gar=0</i>	
	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.
<i>duracao</i>	-0,0050	0,0013 *	-0,0061	0,0003 *	-0,0049	0,0013 *	-0,0061	0,0003 *
<i>relac</i>	-0,0003	0,0003	-0,0002	0,0000 *	-0,0004	0,0002	-0,0002	0,0000 *
<i>risco</i>	0,0065	0,0043	0,0155	0,0013 *	0,0065	0,0042	0,0155	0,0013 *
<i>portecli</i>	-0,0854	0,0049 *	-0,1433	0,0029 *	-0,0850	0,0049 *	-0,1433	0,0028 *
<i>totif</i>	-0,0043	0,0012 *	-0,0114	0,0004 *	-0,0043	0,0012 *	-0,0114	0,0004 *
<i>totoper</i>	-0,1595	0,0465 *	0,0290	0,0083 *	-0,1624	0,0466 *	0,0289	0,0083 *
<i>dvenc</i>	0,0641	0,0058 *	0,0423	0,0031 *	0,0637	0,0059 *	0,0423	0,0031 *
<i>dprej</i>	0,0197	0,0176	-0,0083	0,0135	0,0191	0,0177	-0,0082	0,0135
<i>_cons</i>	4,5483	0,0311 *	4,6178	0,0244 *	4,5467	0,0304 *	4,6191	0,0246 *
<i>sigma</i>	0,1908	0,0152 *	0,1972	0,0047 *	0,1911	0,0153 *	0,1972	0,0047 *
<i>rho</i>	0,2636	0,1310 #	0,0395	0,0248	0,2804	0,1288 &	0,0433	0,0242 #
<i>Dummies:</i>	IFs				IFs			
<i>Restrição exclusão.:</i>	-				<i>cap, ctrabalho2, dummies regiões</i>			

Nº de observações: 51.617. Estimação usando o comando *movestay* (Lokshin e Sajaia, 2004).

Desvio-padrão robusto à heterocedasticidade. Coeficientes das *dummies* de IFs não reportados.

A estatística do teste de Wald para a independência das equações resultou em $Prob > \chi^2 = 0,00$ em ambos os casos.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

2.5.5 Estimações adicionais

A maturidade (*duracao*) também é outra possível variável endógena num contrato de crédito. No entanto, na modalidade analisada (Capital de Giro), devido ao curto prazo de duração do empréstimo (na base que trabalhamos ela é inferior a um ano), cremos que o problema da simultaneidade não seja tão grave. Para encontrar alguma evidência acerca deste fato, estimamos novamente as especificações da **Tabela 2.5**, p.84 e da **Tabela 2.7**, p.88 sem

considerar a variável *duracao*. Os resultados encontram-se no **APÊNDICE 2.C**, p.106 na **Tabela A - 2.4** e na **Tabela A - 2.5**, respectivamente. Verifica-se que a exclusão da *duracao* em geral não afeta substancialmente as estimativas das demais variáveis. A diferença mais significativa refere-se ao coeficiente da variável *risco*, no *probit* com e sem endogeneidade.

Cada cliente da base de dados utilizada (51.617 observações) pode ter uma ou mais operações com uma ou mais IFs. Para diminuir o peso de um cliente que aparece diversas vezes, selecionamos apenas a operação com a maior taxa de juros para cada cliente, reduzindo a amostra a 21.270 observações (sendo 14,51% com garantia). Este valor pode ser pensado como a taxa de juros marginal que o tomador está disposto a pagar por uma unidade de recurso adicional. Estimamos novamente as regressões apresentadas nas subseções **2.5.1** e **2.5.2**, cujas estatísticas descritivas e os resultados encontram-se no **Apêndice 2.D**, p.108 (**Tabela A - 2.6** a **Tabela A - 2.13**).

Na maioria das vezes, os sinais dos coeficientes e a sua significância estão de acordo com os apresentados anteriormente. Na estimação do *probit*, a maior diferença refere-se aos sinais dos coeficientes da *duracao* e do *risco* e a magnitude do coeficiente de *totoper*. *Duracao* é negativo quando se considera a endogeneidade da taxa de juros e *risco* apresenta sinal negativo nas três especificações, sendo não significativo para (2) e (3). O coeficiente da variável *totoper* é menor em módulo na estimação da base reduzida.

Na equação da taxa de juros, a estimação por OLS ou 2SLS apresenta o coeficiente da *duracao* negativo e significativo um número maior de vezes. De modo geral, para os demais coeficientes tanto o sinal quanto a significância são parecidos com o da amostra completa.

Na estimação por *matching*, utilizamos 25% da amostra (=5.318 observações) para tornar o processamento mais rápido. Os coeficientes estimados com a amostra reduzida são em geral, maiores em valores absolutos quando não inclui as *dummies* de IFs e próximos quando se considera essas *dummies*. No efeito do tratamento com endogeneidade, a especificação (2) é a que produziu resultados mais divergentes. Neste caso, o valor do coeficiente da garantia diminui em módulo passando de -0,56 (**Tabela 2.9**, p.93) para -0,23 (**Tabela A - 2.12**, p.111). Ademais, o coeficiente de correlação entre a equação binária e contínua é não significativo.

De todas as técnicas utilizadas, a que mais apresentou divergência no resultado com a redução da amostra foi o *endogenous switching regression model*. Nas duas estimativas apresentadas, o coeficiente *sigma* é significativo sendo um indício de que a mudança de regime é endógena. No entanto, o coeficiente de correlação *rho* é não-significativo em todos os casos indicando que não existe diferença entre o contrato com (e sem) garantia e um contrato aleatório da amostra.

Exceto para a última técnica apresentada, o resultado com a amostra completa e reduzida são parecidos, ou seja, na equação de garantia, a taxa de juros e o tempo de relacionamento têm coeficientes negativos e significativos, o mesmo ocorrendo com as variáveis garantia e relacionamento na equação da taxa de juros.

2.6 CONCLUSÃO

Os dois objetivos deste capítulo são: 1) verificar a relação existente entre garantia e taxas de juros; e 2) verificar qual o impacto do tempo de relacionamento nos termos dos contratos de crédito. O teste empírico foi conduzido com os dados do SCR – Sistema de Informações de Crédito do Banco Central do Brasil (*cross-section* de set/2005). A modalidade escolhida foi Capital de Giro, pois a mesma não possui necessariamente garantias atreladas à operação, como é o caso do financiamento de veículos onde o próprio veículo financiado serve de garantia.

Como não é clara qual a relação de causalidade existente entre garantia e taxa de juros, estimamos duas equações: 1) Inicialmente, consideramos a garantia como variável dependente e a taxa de juros como variável explicativa. Utilizamos o modelo *probit* com e sem endogeneidade para estimar as regressões. 2) Em seguida, invertemos a ordem, utilizando como variável dependente a taxa de juros e a garantia como variável explicativa. Este modelo foi estimado por mínimos quadrados ordinários, variáveis instrumentais, pelo estimador de *matching* proposto por Abadie *et al.* (2004 e 2007) e por efeitos de tratamento com endogeneidade. Em todas as especificações o tempo de relacionamento foi considerado variável exógena. Nas equações binárias, a hipótese adotada é a da normalidade conjunta dos erros.

O primeiro resultado é que em todos os casos encontramos evidências de que na modalidade de crédito estudada, o efeito que predomina é o do paradigma da informação privada. A taxa de juros e a garantia são variáveis substitutas nos contratos de crédito: empréstimos com garantia são aqueles que apresentam baixa taxa de juros. Segundo a teoria de Bester (1985), os bancos ofertam dois tipos de contratos: 1) muita garantia e taxa de juros baixa; e 2) pouca garantia e taxa de juros alta. Os clientes de baixo risco selecionam o primeiro contrato e o cliente arriscado, o segundo contrato. Este resultado está em consonância com o apresentado por Rodrigues *et al.* (2004), utilizando também a base do SCR, para a *cross-section* de dez/2003.

As estimativas mostram indícios de que o efeito que prevalece com um maior tempo de relacionamento é o da redução da assimetria informacional. Um relacionamento mais

duradouro entre o tomador e a IF contribui para a redução da taxa de juros/garantia. Este é o nosso segundo resultado.

Testar a presença de assimetria informacional está além do escopo deste trabalho. No entanto, se os clientes selecionam os diversos contratos baseados em características não-observadas pelo banco, então é possível que contratos com e sem garantia tenham coeficientes distintos na equação da taxa de juros. Neste caso, temos um problema de seleção amostral, o que motivou o uso do *switching regression model*. As estimativas mostram que os empréstimos com garantia possuem seleção negativa, ou seja, eles apresentam características não-observáveis que fazem com que a taxa de juros seja menor relativo a um contrato aleatório da amostra. Os contratos sem garantia, por sua vez, apresentam seleção positiva, resultando em uma taxa de juros maior que um contrato aleatório.

A importância de verificar a relação entre taxa de juros e garantia é ainda maior no contexto brasileiro, onde a recuperação das garantias nem sempre ocorre de maneira rápida devido a diversos problemas relacionados com o judiciário. Se o judiciário é pouco eficiente, a garantia pode não impactar significativamente a determinação da taxa de juros. Apesar de não considerarmos explicitamente o judiciário nas estimativas apresentadas, a conclusão geral é que, apesar da percepção de que o judiciário brasileiro seria pouco eficiente, a garantia é um importante sinalizador da qualidade do tomador.

Como sugestão de pesquisa futura, pode-se testar a existência de assimetria de informação (seleção adversa e risco moral), bem como o efeito do judiciário nas relações contratuais ou um melhor entendimento da relação entre as diversas variáveis que compõem um contrato de crédito e que, muitas vezes, são determinadas de forma simultânea (taxa de juros, garantia, maturidade e valor do empréstimo).

2.7 REFERÊNCIAS BIBLIOGRÁFICAS

ABADIE, Alberto; DRUKKER, David; HERR, Jane Leber; IMBENS, Guido W. *Implementing matching estimators for average treatment effects in Stata*. **The Stata Journal**. 4, Number 3 pp. 290-311. 2004.

ABADIE, Alberto; IMBENS, Guido W. *Bias-corrected matching estimators for average treatment effects*. WP disponível em: <<http://ksghome.harvard.edu/~aabadie/bcm.pdf>>. Acesso em 16/06/2008. Aug., 2007. Versão anterior do paper: *Simple and bias-corrected matching estimators for average treatment effects*. **NBER Technical Working Paper, T0283**. Oct., 2002.

BERGER, Allen N.; UDELL, Gregory F. *Collateral, loan quality and bank risk*. **Journal of Monetary Economics**. 25, pp. 21-42. 1990.

BERGER, Allen N.; UDELL, Gregory F. *Relationship lending and lines of credit in small firm finance*. **The Journal of Business**. Vol. 68, N° 3, pp. 351-381. 1995.

BESTER, Helmut. *Screening vs. rationing in credit markets with imperfect information*. **The American Economic Review**. Vol. 75, N° 75, pp. 850-855. Sep., 1985.

BOOT, Arnoud W. A.; THAKOR, Anjan V. *Moral hazard and secured lending in an infinitely repeated credit market game*. **International Economic Review**. Vol. 35, N° 4, pp. 899-920. Nov., 1994.

BOOT, Arnoud W. A.; THAKOR, Anjan V.; UDELL, Gregory F. *Secured lending and default risk: equilibrium analysis, policy implications and empirical results*. **The Economic Journal**. Vol. 101, N° 406, pp. 458-472. May, 1991.

CAMERON, A. Colin; TRIVEDI, Pravin K. ***Microeconometrics: methods and applications***. Cambridge University Press. 2005.

CHAN, Yuk-Shee; KANATAS, George. *Asymmetric valuations and the role of collateral in loan agreements*. **Journal of Money, Credit and Banking**. Vol. 17, N° 1, pp. 84-95, Feb., 1985.

DEGRYSE, Hans; VAN CAYSEELE, Patrick. *Relationship lending within a bank-based system: evidence from European small business data*. **Journal of Financial Intermediation**. 9, pp. 90-109, 2000.

FREIXAS, Xavier; ROCHET, Jean-Charles. ***Microeconomics of banking***. The MIT Press, 1997.

JIMÉNEZ, Gabriel; SALAS, Vicente; SAURINA, Jesús. *Determinants of collateral*. Documento de Trabajo N° 0420, **Banco de España**. 2004.

JIMÉNEZ, Gabriel; SAURINA, Jesús. *Collateral, type of lender and relationship banking as determinants of credit risk*. **Journal of Banking & Finance**. 28, pp. 2191-2212. 2004.

LOKSHIN, Michael; SAJAIA, Zurab. *Maximum likelihood estimation of endogenous switching regression models*. **The Stata Journal**. 4, Number 3, pp. 282-289. 2004.

MADDALA, G. S. *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press. 1983.

MANOVE, Michael; PADILLA, A. Jorge; PAGANO, Marco. *Collateral versus project screening: a model of lazy banks*. **The RAND Journal of Economics**. Vol. 32, N° 4, pp. 726-44. 2001.

ORTOLANI, Edna Mendes. **Operações de crédito no mercado financeiro**. Editora Atlas. 2000.

ONO, Arito; UESUGI, Ichihiro. *The role of collateral and personal guarantees in relationship lending: evidence from Japan's SME loan market*. Working Paper. Mar, 2008. Disponível em: <http://www.ier.hit-u.ac.jp/pie/stage2/Japanese/discussionpaper/dp2008/dp371/text.pdf>. Acesso em 18/06/2008.

PINHEIRO, Armando Castelar; CABRAL, Célia. Mercado de crédito no Brasil: o papel do judiciário e de outras instituições. **Ensaio BNDES n° 9**. Rio de Janeiro, dez/1998.

RODRIGUES, Eduardo A. de Souza; TAKEDA, Tony; ARAÚJO, Aloísio P. de. Qual o impacto das garantias reais na taxas de juros de empréstimo bancário no Brasil? Uma breve avaliação com base nos Dados do SCR. **Economia Bancária e Crédito - Avaliação de 5 anos do Projeto Juros e Spread Bancário**. Banco Central do Brasil. Dez., 2004.

SHARPE, Steven A. *Asymmetric information, bank lending and implicit contracts: a stylized model of customer relationships*. **The Journal of Finance**. Vol. 45, N° 4, pp. 1069-1087. Sep., 1990.

STOCK, James H.; WATSON, Mark W. *Introduction to econometrics*. Addison Wesley. 2003.

WEILL, Laurent; GODLEWSKI, Christophe. *Does collateral helps mitigate adverse selection? A cross-country analysis*. **MPRA Munich Personal RePEc Archive, Paper N° 2508**. Nov. 2006. Disponível em: http://mpa.ub.uni-muenchen.de/2508/1/MPRA_paper_2508.pdf. Acesso em: 19/02/2008.

WOOLDRIDGE, Jeffrey M. *Econometric analysis of cross section and panel data*. The MIT Press. 2002.

APÊNDICE 2.A

Tabela A - 2.1 – Medidas de eficiência do judiciário por UF - 2005

UF	<i>magist</i>	<i>ctrabalho2</i>	<i>tcongest2</i>
Rondônia	8,21	672,71	40,71
Acre	8,07	269,67	36,88
Amazonas	5,72	745,89	22,23
Roraima	10,23	404,71	57,22
Pará	3,97	385,40	41,04
Amapá	11,45	224,89	29,15
Tocantins	8,05	529,58	78,58
Maranhão	4,33	402,00	30,72
Piauí	4,79	150,12	4,47
Ceará	4,64	6.016,36	94,56
Rio Grande do Norte	8,56	891,40	35,73
Paraíba	7,04	1.034,83	39,96
Pernambuco	5,71	1.908,46	92,62
Alagoas	4,88	348,82	22,13
Sergipe	6,30	1.092,77	45,04
Bahia	4,01	2.909,83	74,29
Minas Gerais	4,76	1.526,27	40,69
Espírito Santo	12,47	1.216,90	40,55
Rio de Janeiro	5,84	875,53	20,38
São Paulo	4,98	2.493,38	58,29
Paraná	6,63	501,62	33,41
Santa Catarina	6,51	1.633,15	51,36
Rio Grande do Sul	7,09	3.211,34	34,96
Mato Grosso do Sul	8,04	1.182,88	30,99
Mato Grosso	9,24	479,74	27,94
Goiás	5,34	1.012,97	34,60
Distrito Federal	11,23	856,97	21,75

Fonte: Conselho Nacional de Justiça.

APÊNDICE 2.B

Tabela A - 2.2 – Resultado da equação de seleção do modelo de efeitos do tratamento

Eq. garantia	(1)		(2)		(3)	
	Coefficiente	Desv. Padr.	Coefficiente	Desv. Padr.	Coefficiente	Desv. Padr.
<i>ctrabalho2</i>	-	-	0,0001	0,0000 *	0,0002	0,0001 *
<i>cap</i>	-	-	0,0045	0,0005 *	-	-
<i>duracao</i>	0,0789	0,0031 *	0,0813	0,0022 *	0,0797	0,0031 *
<i>relac</i>	0,0054	0,0007 *	-0,0092	0,0008 *	0,0056	0,0007 *
<i>risco</i>	0,1675	0,0158 *	0,0404	0,0067 *	0,1685	0,0162 *
<i>portecli</i>	0,4776	0,0204 *	0,0220	0,0210	0,4945	0,0207 *
<i>totif</i>	0,0626	0,0045 *	0,0631	0,0038 *	0,0641	0,0049 *
<i>totoper</i>	-2,1594	0,2589 *	-0,7556	0,1306 *	-2,2969	0,2905 *
<i>dvenc</i>	-0,0574	0,0257 &	-0,0975	0,0183 *	-0,0536	0,0259 &
<i>dprej</i>	0,0416	0,0824	0,2293	0,0635 *	0,0656	0,0820
<i>cons</i>	-0,4538	0,2603 #	-2,3060	0,0554 *	-1,3367	0,4164 *
<i>Dummies</i>	IFs		-		IFs	
<i>Restr. exclus.</i>	-		<i>cap, ctrabalho2</i>		<i>dummies regiões, ctrabalho2</i>	

Nº de observações: 51.617. Desvio-padrão robusto à heterocedasticidade.

Coefficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

A troca de *ctrabalho2* por *magist* ou *tcongest2* não alteram significativamente os resultados.

Tabela A - 2.3 – Resultado da equação de seleção do *endogenous switching regression model*

gar	(1)		(2)	
	Coefficiente	Desvio-Padrão	Coefficiente	Desvio-Padrão
<i>relac</i>	0,0057	0,0007 *	0,0059	0,0007 *
<i>portecli</i>	0,4887	0,0210 *	0,5070	0,0213 *
<i>totif</i>	0,0640	0,0045 *	0,0657	0,0049 *
<i>totoper</i>	-2,2395	0,2623 *	-2,3812	0,2923 *
<i>dvenc</i>	-0,0523	0,0257 &	-0,0495	0,0259 #
<i>dprej</i>	0,0549	0,0818	0,0807	0,0812
<i>duracao</i>	0,0770	0,0032 *	0,0776	0,0032 *
<i>risco</i>	0,1637	0,0160 *	0,1642	0,0163 *
<i>ctrabalho2</i>	-	-	0,0002	0,0001 *
<i>cap</i>	-	-	-0,0026	0,0016
<i>_cons</i>	-0,5147	0,2591 &	-1,3617	0,4156 *
<i>Dummies:</i>	IFs		IFs	
<i>Restr. exclusão.:</i>	-		<i>cap, ctrabalho2, dummies regiões</i>	

Nº de observações: 51.617. Estimação usando o comando *movestay* (Lokshin e Sajaia, 2004).

Desvio-padrão robusto à heterocedasticidade. Coeficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

APÊNDICE 2.C

Abaixo, as estimativas das especificações idênticas às apresentadas ao longo do texto, sem considerar a variável *duracao*.

Tabela A - 2.4 – Resultados das estimativas do modelo *probit* – sem a variável *duração*

Variável	(1)			(2)			(3)		
	Coef.	Desv. Padr.		Coef.	Desv. Padr.		Coef.	Desv. Padr.	
<i>ljuros</i>	-1,5665	0,0312	*	-4,4178	0,0609	*	-4,4596	0,0648	*
<i>relac</i>	-0,0210	0,0010	*	-0,0154	0,0007	*	-0,0163	0,0007	*
<i>risco</i>	0,0911	0,0072	*	0,1194	0,0073	*	0,1155	0,0075	*
<i>portecli</i>	-0,0490	0,0145	*	-0,6470	0,0184	*	-0,5995	0,0206	*
<i>totif</i>	0,0748	0,0052	*	-0,0062	0,0053		-0,0033	0,0058	
<i>totoper</i>	-2,9981	0,3366	*	-1,9195	0,2343	*	-2,1123	0,2701	*
<i>dvenc</i>	0,0739	0,0195	*	0,3218	0,0165	*	0,3157	0,0171	*
<i>dprej</i>	0,1796	0,0614	*	0,2127	0,0591	*	0,1717	0,0582	*
<i>_cons</i>	3,9993	0,2946	*	17,4682	0,2908	*	17,5150	0,3899	*
<i>/athrho</i>	-	-		1,3992	0,0763	*	1,3397	0,0766	*
<i>Dummies:</i>	<i>regiões</i>			-			<i>regiões</i>		
<i>Instrum.:</i>	-			<i>hdep, pot, dummies IFs</i>			<i>hdep, pot, dummies IFs</i>		

A variável dependente é *gar*. N° de observações: 51.564 (1) e (3) / 51.617 (2).

Desvio-padrão robusto à heterocedasticidade. Coeficientes das *dummies* de regiões não reportados.

As colunas (2) e (3) consideram endogeneidade da taxa de juros. Teste de *Wald* para exogeneidade (*/athrho*=0).

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Tabela A - 2.5 – Resultados das estimativas da equação da taxa de juros – OLS e 2SLS – sem a variável duração

Variável	OLS (1)	OLS (2)	2SLS (3)	2SLS (4)	2SLS (5)	2SLS (6)
<i>gar</i>	-0,1757 * (0,0030)	-0,1785 * (0,0039)	-1,7072 * (0,1518)	-0,1488 * (0,0182)	-0,8567 * (0,1990)	-0,2733 * (0,0429)
<i>relac</i>	-0,0021 * (0,0000)	-0,0001 * (0,0000)	-0,0050 * (0,0003)	-0,0020 * (0,0001)	0,0001 (0,0001)	-0,0001 & (0,0000)
<i>risco</i>	-0,0229 * (0,0012)	0,0097 * (0,0012)	0,0051 (0,0036)	-0,0234 * (0,0012)	0,0207 * (0,0037)	0,0112 * (0,0014)
<i>portecli</i>	-0,1710 * (0,0024)	-0,1262 * (0,0022)	-0,1325 * (0,0069)	-0,1717 * (0,0024)	-0,0785 * (0,0144)	-0,1196 * (0,0037)
<i>totif</i>	-0,0118 * (0,0004)	-0,0104 * (0,0003)	0,0096 * (0,0024)	-0,0122 * (0,0005)	-0,0056 * (0,0015)	-0,0097 * (0,0005)
<i>totoper</i>	-0,3024 * (0,0111)	-0,0111 (0,0077)	-0,4992 * (0,0288)	-0,2989 * (0,0112)	-0,0818 * (0,0237)	-0,0210 (0,0087)
<i>dvenc</i>	0,0782 * (0,0032)	0,0467 * (0,0028)	0,0670 * (0,0073)	0,0784 * (0,0032)	0,0413 * (0,0040)	0,0460 * (0,0028)
<i>dprej</i>	0,0108 (0,0122)	-0,0046 (0,0112)	0,1122 * (0,0332)	0,0090 (0,0122)	-0,0002 (0,0144)	-0,0040 (0,0113)
<i>_cons</i>	4,3757 * (0,0057)	4,8046 * (0,0170)	4,4045 * (0,0125)	4,3752 * (0,0057)	5,2899 * (0,1473)	4,8725 * (0,0361)
<i>Dummies:</i>	-	IFs	-	-	IFs	IFs
<i>Instrumentos:</i>	-	-	<i>cap,</i> <i>ctrabalho2</i>	<i>dummies</i> <i>regiões, cap,</i> <i>ctrabalho2</i>	<i>cap,</i> <i>ctrabalho2</i>	<i>dummies</i> <i>regiões, cap,</i> <i>ctrabalho2</i>
<i>Estatística-F do 1º Estágio:</i>	-	-	705,51	127,24	2.877,73	1.452,89
<i>Shea's partial R²</i>			0,0025	0,0325	0,0005	0,0063
<i>Estatística J P-value(chi2)</i>			0,3396	0,0000	0,0001	0,0000

A variável dependente é *ljuros*. N° de observações: 51.617.

Desvio-padrão robusto à heterocedasticidade entre parênteses. Coeficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

APÊNDICE 2.D

Nas tabelas a seguir, apresentamos a estatística descritiva dos dados e os coeficientes estimados quando reduzimos a base a apenas uma única operação por cliente (aquela que possui a maior taxa de juros). As especificações das equações são idênticas às apresentadas ao longo do texto.

Tabela A - 2.6 – Estatística descritiva das variáveis – Capital de Giro – Base de dados reduzida

Variável	Média	Desvio-Padrão	1º quartil	Mediana	3º quartil
<i>juros</i>	54,71	14,42	42,58	54,65	66,69
<i>gar</i>	0,14	0,35	0	0	0
<i>duracao</i>	8,06	4,30	4,00	10,10	12,17
<i>relac</i>	6,68	13,27	0,75	3,21	7,61
<i>risco</i>	2,43	1,11	1	2	3
<i>portecli</i>	2,05	0,57	2	2	2
<i>totif</i>	4,46	2,92	2	4	6
<i>totoper</i>	0,06	0,07	0,03	0,04	0,07
<i>dvenc</i>	0,20	0,40	0	0	0
<i>dprej</i>	0,01	0,11	0	0	0

Nº Observações: 21.270. Fonte: SCR – Bacen (set/2005).

Tabela A - 2.7 – Estatística descritiva das variáveis por garantia – Capital de Giro – Base de dados reduzida

Variável	gar=1 Nº obs: 3.083 (14,5%)		gar=0 Nº obs: 18.187 (85,5%)		Teste de Média $H_0: \mu_1 = \mu_0$
	Média	Desv. Pad.	Média	Desv. Pad.	Estat-Z
<i>juros</i>	41,26	10,43	56,98	13,74	-73,56
<i>duracao</i>	9,06	3,63	7,89	4,38	16,03
<i>relac</i>	3,12	4,75	7,29	14,13	-30,83
<i>risco</i>	2,53	0,82	2,41	1,15	7,04
<i>portecli</i>	2,17	0,85	2,03	0,51	8,88
<i>totif</i>	5,65	3,91	4,26	2,67	19,00
<i>totoper</i>	0,07	0,10	0,05	0,06	10,78
<i>dvenc</i>	0,21	0,41	0,20	0,40	1,26
<i>dprej</i>	0,02	0,14	0,01	0,10	3,80

Fonte: SCR – Bacen (set/2005).

Tabela A - 2.8 – Resultados das estimativas do modelo *probit* – Base de dados reduzida

Variável	(1)		(2)		(3)	
	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.
<i>ljuros</i>	-2,3551	0,0528 *	-4,6835	0,0410 *	-4,6999	0,0407 *
<i>duracao</i>	0,0120	0,0031 *	-0,0090	0,0026 *	-0,0092	0,0026 *
<i>relac</i>	-0,0575	0,0051 *	-0,0123	0,0013 *	-0,0123	0,0013 *
<i>risco</i>	-0,0783	0,0120 *	0,0001	0,0156	-0,0009	0,0155
<i>portecli</i>	-0,1823	0,0256 *	-0,5782	0,0199 *	-0,5818	0,0201 *
<i>totif</i>	0,0568	0,0057 *	-0,0267	0,0049 *	-0,0246	0,0049 *
<i>totoper</i>	-0,7021	0,2233 *	-0,7801	0,2156 *	-0,7875	0,2175 *
<i>dvenc</i>	0,0947	0,0322 *	0,2890	0,0235 *	0,2898	0,0235 *
<i>dprej</i>	0,3901	0,1051 *	0,0690	0,0901	0,0876	0,0890
<i>_cons</i>	7,8038	0,5435 *	19,1100	0,1802 *	19,2144	0,3015 *
<i>/athrho</i>	-	-	1,5353	0,0449 *	1,5454	0,0454 *
<i>Dummies:</i>	<i>regiões</i>		-		<i>regiões</i>	
<i>Instrumentos:</i>	-		<i>hdep, pot, dummies IFs</i>		<i>hdep, pot, dummies IFs</i>	

A variável dependente é *gar*. N° de observações: 21.270 (2) / 21.208 (1) e (3).

Desvio-padrão robusto à heterocedasticidade. Coeficientes das *dummies* de regiões não reportados.

As colunas (2) e (3) consideram endogeneidade da taxa de juros. Teste de *Wald* para exogeneidade (*/athrho*=0).

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Tabela A - 2.9 – Efeitos marginais do modelo *probit* – Base de dados reduzida

Variável	(1)		(2)		(3)	
	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.
<i>ljuros</i>	-0,3245	0,0100 *	-1,2497	0,0226 *	-1,2544	0,0241 *
<i>duracao</i>	0,0016	0,0004 *	-0,0024	0,0007 *	-0,0024	0,0007 *
<i>relac</i>	-0,0079	0,0005 *	-0,0033	0,0003 *	-0,0033	0,0003 *
<i>risco</i>	-0,0108	0,0017 *	0,0000	0,0042	-0,0002	0,0042
<i>portecli</i>	-0,0251	0,0037 *	-0,1543	0,0058 *	-0,1553	0,0061 *
<i>totif</i>	0,0078	0,0008 *	-0,0071	0,0013 *	-0,0066	0,0013 *
<i>totoper</i>	-0,0967	0,0309 *	-0,2082	0,0576 *	-0,2102	0,0581 *
<i>dvenc</i> [@]	0,0136	0,0049 *	0,0829	0,0073 *	0,0831	0,0074 *
<i>dprej</i> [@]	0,0700	0,0234 *	0,0190	0,0255	0,0243	0,0256
<i>Dummies:</i>	<i>regiões</i>		-		<i>regiões</i>	
<i>Instrumentos:</i>	-		<i>hdep, pot, dummies IFs</i>		<i>hdep, pot, dummies IFs</i>	

A variável dependente é *gar*. N° de observações: 21.270 (2) / 21.208 (1) e (3). Efeitos marginais das *dummies* de regiões não reportados. As colunas (2) e (3) consideram endogeneidade da taxa de juros.

[@]: o efeito marginal é calculado para uma variação discreta da *dummy* de 0 para 1.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Tabela A - 2.10 – Resultados das estimativas da equação da taxa de juros – OLS e 2SLS – Base de dados reduzida

Variável	OLS (1)	OLS (2)	2SLS (3)	2SLS (4)	2SLS (5)	2SLS (6)
<i>gar</i>	-0,2707 * 0,0045	-0,1813 * 0,0059	-1,5531 * 0,2130	-0,5828 * 0,0374	-0,4294 * 0,1602	-0,3220 * 0,0712
<i>duracao</i>	-0,0055 * 0,0004	-0,0082 * 0,0004	0,0032 & 0,0016	-0,0034 * 0,0005	-0,0061 * 0,0014	-0,0070 * 0,0007
<i>relac</i>	-0,0001 0,0001	-0,0008 * 0,0001	-0,0040 * 0,0007	-0,0011 * 0,0002	-0,0008 * 0,0001	-0,0008 * 0,0001
<i>risco</i>	-0,0473 * 0,0015	0,0092 * 0,0023	-0,0478 * 0,0026	-0,0474 * 0,0016	0,0109 * 0,0027	0,0102 * 0,0024
<i>portecli</i>	-0,1280 * 0,0039	-0,1340 * 0,0037	-0,1020 * 0,0099	-0,1217 * 0,0045	-0,1163 * 0,0120	-0,1239 * 0,0062
<i>totif</i>	-0,0119 * 0,0010	-0,0093 * 0,0009	0,0144 * 0,0049	-0,0055 * 0,0014	-0,0071 * 0,0017	-0,0080 * 0,0011
<i>totoper</i>	-0,1739 * 0,0466	-0,1051 & 0,0422	-0,2564 * 0,0940	-0,1940 * 0,0544	-0,1074 & 0,0439	-0,1064 & 0,0428
<i>dvenc</i>	0,0740 * 0,0043	0,0548 * 0,0040	0,0336 * 0,0112	0,0641 * 0,0049	0,0512 * 0,0048	0,0527 * 0,0042
<i>dprej</i>	-0,0137 0,0177	-0,0193 0,0165	0,1015 & 0,0458	0,0143 0,0207	-0,0061 0,0199	-0,0118 0,0176
<i>_cons</i>	4,4735 * 0,0085	4,8529 * 0,0366	4,4569 * 0,0192	4,4695 * 0,0096	5,0134 * 0,1138	4,9440 * 0,0623
<i>Dummies:</i>	-	IFs	-	-	IFs	IFs
<i>Instrumentos:</i>	-	-	<i>cap,</i> <i>ctrabalho2</i>	<i>dummies</i> <i>regiões, cap,</i> <i>ctrabalho2</i>	<i>cap,</i> <i>ctrabalho2</i>	<i>dummies</i> <i>regiões, cap,</i> <i>ctrabalho2</i>
<i>Estatística-F do 1º Estágio:</i>	-	-	122,54	32,91	1.029,82	534,50
<i>Shea's partial R²</i>	-	-	0,0020	0,0185	0,0013	0,0062
<i>Estatística J P-value(chi2)</i>	-	-	0,0043	0,0000	0,0000	0,0000

A variável dependente é *ljuros*. N° de observações: 21.270.

Desvio-padrão robusto à heterocedasticidade entre parênteses. Coeficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Tabela A - 2.11 Resultados das estimativas do matching estimator – Base de dados reduzida.

Estimação		ATE	ATT	ATC
Coeficiente	(1)	-0,2013 *	-0,1646 *	-0,2072 *
Desvio-padrão		(0,0189)	(0,0119)	(0,0211)
Coeficiente	(2)	-0,1549 *	-0,0442 *	-0,1724 *
Desvio-padrão		(0,0498)	(0,0146)	(0,0572)

A variável dependente é *ljuros*. N° de observações: 25% da amostra=5.318.

Estimação usando o comando *nnmatch* (Abadie *et al.*, 2004).

(2) inclui *dummies* de IFs. Desvio-padrão com correção de viés.

* variável associada significativa a 1%.

Tabela A - 2.12 – Resultados das estimativas do modelo de efeitos do tratamento – Base de dados reduzida

Eq. ln(juros)	(1)		(2)		(3)	
	Coeficiente	Desv. Padr.	Coeficiente	Desv. Padr.	Coeficiente	Desv. Padr.
<i>gar</i>	-0,2610	0,0206 *	-0,2331	0,0299 *	-0,2713	0,0197 *
<i>duracao</i>	-0,0075	0,0004 *	-0,0058	0,0004 *	-0,0074	0,0004 *
<i>relac</i>	-0,0008	0,0001 *	0,0000	0,0002	-0,0008	0,0001 *
<i>risco</i>	0,0097	0,0023 *	-0,0473	0,0015 *	0,0098	0,0023 *
<i>portecli</i>	-0,1283	0,0040 *	-0,1288	0,0040 *	-0,1275	0,0040 *
<i>totif</i>	-0,0086	0,0009 *	-0,0126	0,0011 *	-0,0085	0,0009 *
<i>totoper</i>	-0,1059	0,0424 &	-0,1715	0,0460 *	-0,1060	0,0425 &
<i>dvenc</i>	0,0536	0,0040 *	0,0751	0,0044 *	0,0535	0,0040 *
<i>dprej</i>	-0,0151	0,0168	-0,0171	0,0178	-0,0145	0,0168
<i>cons</i>	4,9045	0,0414 *	4,4740	0,0086 *	4,9112	0,0417 *
<i>rho</i>	0,2216	0,0544 *	-0,0886	0,0687	0,2539	0,0524 *
<i>sigma</i>	0,2150	0,0014 *	0,2398	0,0015 *	0,2153	0,0014 *
<i>Dummies</i>	IFs		-		IFs	
<i>Restrição de exclusão:</i>	-		<i>cap, ctrabalho2</i>		<i>dummies regiões, ctrabalho2</i>	

Nº de observações: 21.270. Desvio-padrão robusto à heterocedasticidade.

Coeficientes das *dummies* de IFs não reportados.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Tabela A - 2.13 – Resultado do *endogenous switching regression model* – Base de dados reduzida

<i>ljuros</i>	(1)				(2)			
	<i>gar=1</i>		<i>gar=0</i>		<i>gar=1</i>		<i>gar=0</i>	
	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.
<i>duracao</i>	-0,0086	0,0018 *	-0,0069	0,0008 *	-0,0085	0,0018 *	-0,0070	0,0007 *
<i>relac</i>	-0,0026	0,0014 #	-0,0007	0,0001 *	-0,0027	0,0014 &	-0,0007	0,0001 *
<i>risco</i>	0,0098	0,0053 #	0,0093	0,0026 *	0,0099	0,0053 #	0,0093	0,0026 *
<i>portecli</i>	-0,0918	0,0066 *	-0,1409	0,0087 *	-0,0918	0,0065 *	-0,1411	0,0078 *
<i>totif</i>	-0,0040	0,0017 &	-0,0095	0,0012 *	-0,0041	0,0016 &	-0,0095	0,0012 *
<i>totoper</i>	-0,0685	0,0424	-0,1107	0,0548 #	-0,0685	0,0424	-0,1104	0,0549 #
<i>dvenc</i>	0,0631	0,0098 *	0,0509	0,0044 *	0,0627	0,0098 *	0,0510	0,0044 *
<i>dprej</i>	-0,0114	0,0306	-0,0138	0,0197	-0,0112	0,0306	-0,0141	0,0195
<i>_cons</i>	4,5763	0,0452 *	4,7372	0,1157 *	4,5759	0,0441 *	4,7307	0,1037 *
<i>sigma</i>	0,1981	0,0202 *	0,2168	0,0107 *	0,1982	0,0201 *	0,2166	0,0099 *
<i>rho</i>	0,2072	0,1468	0,4157	0,2900	0,2156	0,1412	0,4091	0,2478
<i>Dummies:</i>	IFs				IFs			
<i>Restrição de exclusão:</i>	-				<i>cap, ctrabalho2, dummies regiões</i>			

Nº de observações: 21.270. Estimação usando o comando *movestay* (Lokshin e Sajaia, 2004).

Desvio-padrão robusto à heterocedasticidade. Coeficientes das *dummies* de IFs não reportados.

A estatística do teste de *Wald* para a independência das equações resultou em *Prob>chi2*=0,0458 e 0,0283 nas especificações (1) e (2), respectivamente.

*, & e # indicam coeficiente da variável associada significativos a 1%, 5% e 10%, respectivamente.

Capítulo 3

3 Modelo de submercados independentes aplicado ao setor bancário brasileiro.

3.1 INTRODUÇÃO

Este estudo tem como objetivo contribuir para o entendimento da estrutura do mercado bancário no Brasil. Em particular, testaremos as previsões do modelo de submercados independentes desenvolvido por Sutton (1998) utilizando dados de agências bancárias de dez/2002. Sob os pressupostos do modelo, o equilíbrio para o agregado de submercados deve ocorrer com um grau mínimo de assimetria na distribuição do tamanho das firmas.

Quanto mais concentrado, maior é a probabilidade das firmas atuantes exercerem poder de mercado (formando um cartel, por exemplo). O bem-estar do consumidor é menor quanto maior o poder de mercado das firmas. E uma das funções dos órgãos de regulação e defesa da concorrência é atuar de forma a diminuir a concentração de mercado, coibindo os abusos causados por uma estrutura de mercado não-competitiva.

Apresentamos alguns dados para ilustrar como tem evoluído a concentração no setor bancário. Na **Tabela 3.1** exibimos o número de bancos (comerciais, múltiplos e a Caixa Econômica) e na **Tabela 3.2** a participação percentual nos ativos totais classificadas pelo porte⁴⁴. O número

⁴⁴ O porte grande, médio e pequeno é definido da seguinte maneira: 1) Relaciona-se a participação relativa do ativo total ajustado de cada instituição com a soma dos ativos totais ajustados de todas as IFs consideradas. As instituições cujo percentual de participação individual é superior a 15% são consideradas de grande porte e excluídas da amostra; 2) Toma-se a amostra dos demais bancos, classifica-se em ordem decrescente das suas participações individuais e calcula-se o acumulado dessas participações; 3) Faz-se os cortes nos acumulados de 70%, 95% e 100%; 4) As IFs que compõem até 70%, inclusive, também são consideradas de grande porte. As IFs entre 70% até 95% são de médio porte e as restantes de pequeno porte. Esta metodologia foi desenvolvida para uso em relatório de evolução do SFN (dos quais retiramos os dados), não se tratando de critério institucional do Banco Central do Brasil.

de instituições financeiras classificadas como grande vem caindo bastante: em 1996 eram contabilizados 28 e em 2006 somente 8 se enquadravam nessa categoria, uma queda de 71,4%. No entanto, a participação nos ativos totais desse segmento não diminuiu no mesmo ritmo. Entre 1996 e 2006, a redução na participação nos ativos totais dos bancos de grande porte foi de apenas -14,4% (de 78,3% para 67,0%). A quantidade de IFs classificadas como de médio porte também se reduziu, passando de 62 em 1996 para 40 em 2006 (-35,5%). Por sua vez, no mesmo período, a participação percentual desse grupo no ativo total aumentou 54,1%.

Tabela 3.1 – Número de instituições do segmento bancário segundo porte por ativos totais - 1996-2006

Porte	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Grande	28	24	25	23	21	17	16	14	13	12	8
Médio	62	63	55	55	56	54	47	49	48	45	40
Pequeno	141	130	123	116	115	111	104	101	103	104	111
Total	231	217	203	194	192	182	167	164	164	161	159

Fonte: Cosif – Banco Central do Brasil.

Em se tratando de bancos de porte pequeno, verifica-se que após uma queda inicial no número de bancos, houve uma certa estabilização. Por outro lado, a participação no percentual de ativos totais apresentou um ligeiro aumento ao longo do período (1996-2006).

Tabela 3.2 – Participação percentual nos ativos totais do segmento bancário - 1996-2006

Porte	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Grande	78,3	79,1	79,7	79,2	79,2	74,0	74,8	74,3	75,2	68,6	67,0
Médio	18,1	18,1	16,9	17,4	17,4	21,8	21,0	21,5	20,7	26,3	27,9
Pequeno	3,6	2,8	3,4	3,4	3,4	4,2	4,2	4,2	4,1	5,1	5,1
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Fonte: Cosif – Banco Central do Brasil.

De modo geral, o fato de o número de bancos estar decrescendo ano a ano, passando de 231 em 1996 para 159 em 2006, pode ser um indício de que este setor tende a se concentrar. Para buscar mais evidências acerca desse fato, na **Tabela 3.3** apresentamos alguns índices de concentração⁴⁵ dos conglomerados bancários para o período de dez/94 a dez/06:

⁴⁵ O índice C_m , onde m é um número inteiro positivo, é comumente utilizado para medir a concentração no mercado e equivale à soma das participações das m maiores firmas no mercado.

Tabela 3.3 – Índices de Concentração (C3, C5 e C10) - 1994-2006 – conglomerados bancários

data	ativo			depósito			crédito		
	C3	C5	C10	C3	C5	C10	C3	C5	C10
dez/94	0,40	0,50	0,65	0,40	0,54	0,69	0,42	0,55	0,72
dez/95	0,37	0,50	0,64	0,44	0,57	0,70	0,42	0,55	0,69
dez/96	0,37	0,49	0,61	0,43	0,57	0,69	0,38	0,52	0,65
dez/97	0,42	0,53	0,66	0,49	0,63	0,74	0,33	0,44	0,60
dez/98	0,44	0,57	0,72	0,54	0,68	0,79	0,37	0,49	0,66
dez/99	0,42	0,55	0,72	0,56	0,69	0,81	0,34	0,48	0,69
dez/00	0,44	0,59	0,78	0,56	0,70	0,84	0,37	0,52	0,72
dez/01	0,44	0,58	0,77	0,54	0,68	0,83	0,40	0,55	0,75
dez/02	0,48	0,63	0,81	0,55	0,70	0,87	0,42	0,58	0,78
dez/03	0,54	0,67	0,85	0,55	0,72	0,90	0,45	0,64	0,83
dez/04	0,51	0,67	0,86	0,53	0,71	0,90	0,45	0,65	0,85
dez/05	0,49	0,65	0,85	0,52	0,70	0,91	0,46	0,67	0,87
dez/06	0,52	0,68	0,89	0,53	0,71	0,92	0,48	0,68	0,88
Média	0,45	0,59	0,76	0,51	0,66	0,82	0,41	0,56	0,75

C_m , $m = 3,5,10$ refere-se à soma das participações do mercado das m maiores firmas.

Fonte dos dados: Banco Central do Brasil.

Nas três primeiras colunas da **Tabela 3.3**, os índices **C3**, **C5** e **C10** são calculados para o ativo. Nas colunas subseqüentes apresentamos os mesmos indicadores para depósito e crédito. O valor **C3**=0,40 de dez/94 na coluna dos ativos indica que as três maiores instituições financeiras da amostra possuem 40% do total de ativos do mercado. Leitura análoga é feita para os demais valores. Em média, no mercado de crédito, as três e as cinco maiores empresas possuem um percentual menor do total se comparado com o depósito ou o ativo. De modo geral, dez bancos representam mais de 60% do mercado, chegando a 92% no caso de depósitos em dez/06. Nos últimos anos esses índices vêm aumentando, indicando tendência à concentração desses mercados.

A análise isolada dos dados brasileiros sugere que o mercado está mais concentrado. No entanto, quando comparamos o Brasil com os demais países, verifica-se que relativamente essa concentração não é muito alta. Na **Figura 3.1**, reproduzimos parte do trabalho desenvolvido por Beck *et al.* (1999) e apresentamos o valor do índice **C3** calculado para os ativos totais:

Figura 3.1 – Concentração dos três maiores bancos (C3) – Média de 1995-99

Fonte dos dados: BankScope, in: Beck *et al.* (2004), reprodução parcial dos dados.

No trabalho de Beck *et al.* (1999) os autores calculam esse índice para um total de 74 países, dentre os quais reproduzimos 19, além da média. O C3 do Brasil é igual a 40%, valor similar ao apresentado na primeira coluna da **Tabela 3.3**, para o mesmo período, e o da média dos 74 países é 61%. Ordenando de forma crescente esse índice para os 74 países, o Brasil encontra-se na 14ª posição, juntamente com o Reino Unido. Existem portanto 59 países que possuem índice de concentração superior, se consideramos a participação dos ativos totais das três maiores instituições financeiras. Outra informação que extraímos da **Figura 3.1** é que tanto países desenvolvidos quanto em desenvolvimento apresentam valores altos do C3 e, portanto, parece não existir uma relação clara entre essas variáveis.

Dessa forma, apesar de o mercado bancário brasileiro apresentar tendência a se concentrar, se comparado com os indicadores de outros países, verifica-se que a concentração não é tão alta. O tema da concentração bancária tem despertado contínuo interesse entre os estudiosos. Nakane (2001) testa o poder de mercado dos bancos brasileiros utilizando a metodologia proposta por Bresnahan (1982) e Lau (1982) e encontra evidências de que os bancos possuem algum poder de mercado; no entanto rejeita-se a hipótese de que atuem segundo um cartel.

Através da estimação por dados em painel, Belaisch (2003) encontra evidências de presença de estrutura de mercado não-competitiva no sistema bancário brasileiro, rejeitando também a hipótese de cartel. Esse fator explicaria porque a intermediação financeira no país é pouco volumosa e custosa. Rocha (2001) apresenta um estudo sobre a evolução da concentração bancária brasileira desde o Plano Real até o ano 2000, utilizando as razões de concentração e o Índice de *Hirschman-Herfindahl* para diversos indicadores (ativos totais, depósitos totais, entre outros). O autor encontra evidências de que existe elevação da concentração no período analisado. O objetivo do trabalho de Cardoso (2004) é verificar se a estrutura de mercado no qual existem firmas dominantes e seguidoras é a que melhor se adapta ao mercado de crédito brasileiro. A autora rejeita essa tese; no entanto, os resultados empíricos suportam a hipótese de uma estrutura assimétrica.

Este trabalho, por sua vez, não pretende testar se a estrutura do mercado bancário brasileiro é melhor representada por um oligopólio, um cartel ou competição perfeita. Nossa pretensão é analisar a estrutura desse mercado sob a luz do modelo de submercados independentes desenvolvido por Sutton (1998). O autor mostra que se pode inferir sobre a distribuição das firmas quando os dados de diversos submercados são agregados e, sob certas condições, é possível derivar o limite inferior dessa distribuição.

Por isso, a assimetria no tamanho das firmas no mercado nacional pode advir da agregação de vários submercados independentes. E um banco pode ser grande no *ranking* nacional porque está presente na maioria dos submercados. Isso não significa necessariamente que em cada submercado, os bancos não competem entre si.

Nesse contexto, o grau de competição entre as firmas deveria ser modelado e testado no nível em que o mesmo ocorre. Por exemplo, uma teoria que descreve o comportamento das firmas no mercado estadual, se testada com dados estaduais pode levar a conclusões erradas se esse mercado for composto por diversos submercados.

O modelo de Sutton (1998), apesar de detectar certa regularidade nos padrões da distribuição no nível agregado, nada prevê (e nem era essa sua intenção) acerca do comportamento esperado das firmas nos submercados nem sobre como definir um mercado relevante. No entanto, a principal lição deste modelo é que o cálculo dos índices de concentração (como os apresentados nessa seção) deve levar em conta se o mercado de interesse é composto por

diversos submercados independentes. Se isso for verdade, as autoridades antitruste não deveriam negligenciar os efeitos da independência na hora de avaliar a estrutura de mercado.

Os principais resultados obtidos neste capítulo são: 1) os resultados para o mercado de agências bancárias é consistente com a teoria de Sutton; 2) os mesmos indícios não são obtidos no mercado de depósito e crédito; e 3) à medida que se agregam os mercados em mesorregiões, UFs e todo o território nacional, o grau de desigualdade dos bancos torna-se cada vez maior.

3.2 MODELO DE SUBMERCADOS INDEPENDENTES

Por que um pequeno número de firmas domina algumas indústrias? Quais as razões que levam à distribuição assimétrica no tamanho das firmas? A resposta a estas perguntas é a motivação de Sutton (1998 e 2007) para o desenvolvimento do modelo de submercados independentes. Inicialmente apresentamos a teoria proposta por Sutton (1998) e comentamos brevemente acerca da definição do mercado. Após, apresentamos alguns trabalhos empíricos que testam as previsões deste modelo e, por fim, discutimos como este modelo poderia ser aplicado ao mercado bancário brasileiro.

3.2.1 Modelo de Submercados Independentes

Esta subseção tem o objetivo de apresentar um resumo sintético do **Modelo de Submercados Independentes** proposto por Sutton (1998).

Dois conceitos são centrais nesta teoria: 1) a definição do mercado e; 2) utilização da abordagem de limites (*bounds approach*). O primeiro será abordado na **subseção 3.2.2**. O segundo decorre do desejo de que o modelo proposto não seja restrito a uma indústria específica, mas que abarque uma grande quantidade delas. Em geral, os resultados empíricos de teoria dos jogos aplicados à organização industrial diferem de modo considerável quando se mudam os pressupostos. Por exemplo, a utilização de um modelo de entrada seqüencial ao invés de simultâneo bem como a adoção de competição de preços ou de quantidades tende a alterar significativamente as conclusões. Neste caso, a abordagem de limites⁴⁶ suporta diferentes modelos e, portanto, se mostra adequada para finalidades empíricas.

⁴⁶ Num espaço possível de realizações, o equilíbrio num modelo de oferta e demanda é determinado por um único ponto. Este resultado, por sua vez, depende de um conjunto de parâmetros específicos da indústria. Para determinar o equilíbrio de forma mais abrangente e que não dependa exclusivamente da forma como o jogo é modelado, o autor utiliza o conceito de limites. Nesta abordagem, o espaço dos resultados é dividido em duas partes, uma contendo os pontos que podem ser suportados como um equilíbrio e outra com os pontos que não podem. O conjunto desses possíveis equilíbrios é determinado através de um jogo por etapas (*stage game approach*).

Sucedese então que o equilíbrio neste modelo ocorre com um grau mínimo de desigualdade no tamanho das firmas, denominado de limite inferior (*lower bound*), cujo conceito está apoiado na definição de mercado.

Um submercado surge quando a cadeia de substituição dos produtos é quebrada. Se um mercado for definido de forma muito ampla é possível que ele seja composto por diversos submercados. Dentro de seus limites, os produtos apresentam alto grau de substituição; e, entre eles, baixo grau de substituição. Um grupo de produtos ou uma diferente localização geográfica podem definir um submercado.

Os submercados são independentes no sentido de que a ligação entre eles é fraca, tanto do lado da oferta (tecnologias de produção) quanto em termos de demanda (substituição dos produtos), de tal modo que a estratégia das firmas em um submercado não é função do que acontece nos demais⁴⁷. Ademais, todas as interações estratégicas ocorrem dentro dos submercados e não entre eles.

O equilíbrio depende de dois estágios de decisões tomadas pela firma. No primeiro, decide-se acerca dos custos afundados para criar o novo produto, instalar uma nova planta ou o gasto com publicidade. No segundo, decide-se acerca dos preços ou das quantidades. Este pressuposto simplificador capta a idéia de que as decisões de investimento do passado afetam as condições de competitividade e lucratividade atuais e futuras, sendo possível fazer afirmações acerca do equilíbrio sem relacioná-lo diretamente com o processo de entrada. O autor então utiliza o conceito técnico de *configuração de equilíbrio* que é definido diretamente no espaço dos resultados (resultados estes que dependem de decisões anteriores) em vez de ser definido no espaço das estratégias (e que levam a determinado resultado).

Formalmente, uma *configuração de equilíbrio*⁴⁸ de produtos é qualquer configuração que satisfaz as seguintes condições:

⁴⁷ Na ausência de despesas com Pesquisa & Desenvolvimento ou com Publicidade & Propaganda que sejam comuns, pode-se considerar que submercados regionais distintos são aproximadamente independentes.

⁴⁸ A *configuração de equilíbrio* é um conceito mais amplo que o *equilíbrio de Nash*, pois este último requer também a condição de ótimo (que por sua vez, depende da forma específica do processo de entrada).

- 1) **Viabilidade** (princípio da sobrevivência): o conjunto de produtos de cada firma cobre seus custos irrecuperáveis; e
- 2) **Estabilidade** (princípio da arbitragem): não existe nenhuma brecha na configuração dos produtos que um entrante possa preencher de maneira lucrativa.

A condição **1)** é garantida pois todas as firmas têm a opção de “não entrar”. A condição **2)** estabelece que não existe desvio lucrativo para a firma.

O pressuposto inicial para compreender a distribuição dos diferentes tamanhos das firmas é que cada submercado é composto por uma única “ilha” que comporta uma planta. Ao longo do tempo, esta ilha se torna acessível diversas vezes às demais firmas presentes no mercado. A idéia é que cada uma das firmas ativas tem a mesma probabilidade de entrar em uma “ilha”, independente do seu tamanho e do seu histórico de entradas em outros submercados.

Este tipo de abordagem permite a modelagem através da teoria dos jogos, onde cada “ilha” pode suportar diversas firmas com diversos tipos de interações estratégicas. Por exemplo, um equilíbrio pode ocorrer com uma firma entrando e todas as demais permanecendo fora do submercado. Outro possível equilíbrio do jogo ocorre quando existe uma seqüência de firmas em um intervalo de tempo discreto que entram em um mercado, cada qual com um número diferente de produtos.

A diferença no tamanho das firmas surge porque uma vez que a primeira firma decide entrar num submercado, então automaticamente ela será diferente das demais firmas que deixaram de entrar. Ademais, se existem muitos submercados e nenhuma firma é grande o suficiente para estar presente na maioria deles, cada firma terá várias oportunidades de investimento (inauguração de nova planta, introdução de novo produto, etc.), tanto nos submercados em que já atua quanto naqueles onde ainda não está presente. Neste caso, todas as firmas que estão fora do submercado têm a mesma probabilidade de entrada.

Neste contexto, as previsões de equilíbrio envolvem assimetria no porte das firmas quando se agregam diversos submercados independentes, mas que não é observada dentro de cada submercado. Estas predições acerca da distribuição do tamanho das firmas são representadas

pela Curva de *Lorenz*, cuja derivação da curva-limite utiliza as condições **1)** e **2)** acima, juntamente com o seguinte princípio:

- 3) **Princípio da Simetria:** Se duas firmas ainda não entraram no submercado s , então ambas possuem estratégias que induzem a estratégias idênticas para o submercado s .

Ou seja, dentro de um submercado, todos os potenciais entrantes recebem o mesmo tratamento independente do seu tamanho. Dessa forma, as oportunidades disponíveis determinarão as estratégias da firma em um submercado e não terão relação com o seu histórico de atuação passado.

Sob estes pressupostos, para qualquer razão fixa $\frac{k}{N}$, o limite inferior assintótico para a concentração das k maiores firmas é dada pela curva de referência:

Equação 3.1:
$$C_k \geq \frac{k}{N} \left(1 - \ln \left(\frac{k}{N} \right) \right)$$

onde N é o número total de firmas no mercado; e

C_k é a razão de concentração de k firmas.

Segue-se então que se os produtos são substitutos próximos e a competição de preços é fraca, as previsões empíricas do modelo são: 1) a curva de *Lorenz* para os submercados individuais deve estar próxima à diagonal; e 2) a curva de *Lorenz* para o agregado de submercados deve estar na posição acima (ou igual) a curva de referência.

3.2.2 Definição do mercado

O primeiro passo para se estudar a concentração e/ou a intensidade de competição entre as firmas é definir qual é o mercado relevante do produto de interesse. A definição de mercado relevante pode considerar a dimensão do produto e a geográfica. Os produtos bancários podem ser os depósitos à vista, a prazo e de poupança e cada uma das diversas modalidades

de empréstimo. A extensão geográfica é a mais difícil de ser delimitada. Ainda que seja possível realizar a maior parte das transações via Internet ou telefone, é mais provável que as pessoas tenham suas contas na região geográfica na qual residem ou trabalham. Neste caso, é factível que o mercado de agências bancárias seja essencialmente geográfico.

Os centros urbanos tendem a atrair trabalhadores que moram em uma região, mas que trabalham em outras e, conseqüentemente, demandam produtos e serviços bancários próximos aos locais que habitualmente freqüentam. Esse fator torna mais difícil a definição do mercado geográfico. Dada a dificuldades de definir claramente os limites do mercado numa área urbana, Bresnahan e Reiss, 1990 (concessionária de veículos) e 1991 (dentistas, médicos e pequenos varejistas) e Abraham *et al.*, 2005 (serviços hospitalares) focam em mercados geograficamente isolados. Definindo dessa forma, esses autores tentam assegurar que: 1) não haja competidores relevantes advindos de outros mercados; e 2) as firmas presentes no mercado compitam umas com as outras.

Um conceito utilizado para se definir um mercado relevante é o da elasticidade de substituição: se dois produtos apresentam elasticidades cruzadas elevadas, então ambos pertencem a um mesmo mercado. Para se garantir que os produtos fora do mercado não sejam substitutos daqueles que estão dentro, uma solução é defini-los de maneira bastante ampla. Neste caso, os produtos que estão dentro do mercado podem ter pouco em comum. Um mercado definido deste modo pode abranger diversos submercados e este é o tema central que motiva a teoria proposta por Sutton (1998).

Empiricamente, os limites administrativos têm sido utilizados para a definição do mercado geográfico. A literatura bancária norte-americana tem utilizado os condados e os denominados *metropolitan statistical areas* (MSAs) que compreendem núcleos populacionais e comunidades adjacentes englobando um ou mais condados. Cetorelli (2002) utiliza os *rural counties* e os MSAs. Cohen e Mazzeo (2004) definem como mercado os *non-MSA labor market areas* (LMAs). Os MSAs *banking markets* são utilizados por Dick (2007) e Zhou (2007) utiliza dados dos condados. No Brasil, os dados de municípios são utilizados por Nakane *et al.*, 2006 (municípios em que havia pelo menos um banco em operação) e Coelho *et al.*, 2007 (municípios em que havia pelo menos um banco em operação, excluídas as regiões metropolitanas).

3.2.3 Alguns resultados empíricos

Sutton (1998) testa a sua teoria no mercado de cimento americano. Considerando todo o mercado, as firmas apresentam tamanhos heterogêneos. Dado que o transporte economicamente viável deste produto ocorre dentro de um estado, este foi identificado como sendo um submercado. Neste nível, a distribuição do tamanho das firmas torna-se mais homogênea. Segundo o autor, uma possível explicação para este fenômeno é que as maiores empresas operam em diversos submercados.

De Juan (2003) verifica as predições deste modelo para o mercado bancário espanhol, utilizando como razão de concentração a proporção de agências bancárias. Os resultados encontrados estão em consonância com a teoria. A autora encontra evidências de que cerca de 67% das cidades com 1000 a 5000 habitantes são considerados submercados independentes e 96,3% deles são totalmente fragmentados. No entanto, quando se considera o mercado nacional ou regional, o grau de desigualdade das firmas se torna elevado. Nosso trabalho se assemelha a este em diversos aspectos.

O mercado bancário da Itália é o tema de estudo de Giannetti (2008). As províncias italianas são definidas como mercado geográfico. Na primeira parte, a autora desenvolve um modelo teórico de acordo com a teoria desenvolvida por Sutton. Na segunda, estima a relação entre a estrutura e o tamanho do mercado, tendo como variável dependente a razão de concentração de uma firma e variáveis relacionadas ao tamanho do mercado como explicativas. A autora encontra evidência sobre a relação inversa entre tamanho e estrutura de mercado.

Os limites administrativos determinam os submercados no mercado de seguro de veículos italiano (Buzzacchi e Valletti, 2006). Na Itália existe uma peculiaridade legal em que os seguros dos veículos são compulsórios e devem ser feitos na província em que o mesmo é registrado, constituindo uma indicação do tamanho mínimo dos submercados. Os autores encontram evidências de que as predições do modelo de Sutton são válidas: no nível dos submercados, as firmas apresentam tamanho menos heterogêneo do que quando se considera o mercado italiano como um todo.

Walsh e Whelan (2002) também testam as previsões do modelo de Sutton, mas ao contrário dos autores anteriormente mencionados, a dimensão do produto é utilizada para definir os mercados relevantes no mercado de refrigerantes irlandeses. Neste mercado também é possível inferir a respeito da distribuição das firmas, que está de acordo com as previsões de Sutton.

3.2.4 Modelo aplicado ao mercado bancário

Nesta subseção discutimos a aplicação do modelo de Sutton (1998) ao mercado bancário. O mercado relevante para o setor bancário pode ser definido na dimensão do produto (depósitos, crédito, etc.) e na dimensão geográfica (localização das agências, etc.). Se considerarmos o mercado de produtos isoladamente, as estratégias dos bancos podem estar integradas a nível nacional ou regional. Por exemplo, um banco pode estar determinado a expandir a carteira de crédito ao consumidor, investindo em publicidade em rede nacional para atingir seu objetivo. Nesta dimensão, não é possível sustentar que os submercados que emergem nessa situação sejam completamente independentes. Por outro lado, se partirmos do pressuposto de que para um consumidor mineiro, os serviços ofertados por uma agência localizada em Belo Horizonte são diferentes daqueles ofertados por outra agência localizada em Recife, então a localização geográfica é importante no mercado bancário⁴⁹. Neste caso, os submercados surgem devido à existência de diferentes locais que um banco pode escolher para sua agência operar dentro do país. Sob o ponto de vista da demanda, um submercado é distinto de outro se a elasticidade cruzada entre eles for igual ou próxima a zero. Por parte da oferta, consideramos que as estratégias dos bancos são tomadas dentro de cada submercado e são independentes entre eles. O mercado bancário nacional consiste na junção de todos eles.

Os dois estágios de decisão da firma são: inicialmente, o banco decide sobre a abertura das agências (custos afundados) e, depois, fixa os preços. O conceito de configuração de equilíbrio é utilizado para encontrar o número de bancos que atuam em cada submercado.

⁴⁹ Argumento análogo pode ser feito dentro de um mesmo município. Se o consumidor prefere os serviços ofertados por uma agência localizada ao lado da sua residência, então outra agência, no extremo oposto da cidade, oferta serviços distintos e, portanto, a localização geográfica seria importante no mercado de agências bancárias.

Sob esses pressupostos, o equilíbrio deve envolver um grau mínimo de assimetria no tamanho dos bancos. A curva-limite (**Equação 3.1**, p.121) é derivada para o agregado de submercados (sob os pressupostos apresentados na **subseção 3.2.1**) e fornece o limite assintótico inferior para a concentração.

No entanto, o modelo também prevê que se os produtos ofertados pelas agências bancárias são substitutos próximos e a competição de preços não é elevada⁵⁰, a curva de Lorenz para os submercados individuais deve estar próxima à diagonal (curva representativa do mercado totalmente fragmentado). Na parte empírica utilizaremos os dados de agências bancárias e, adicionalmente, faremos alguma análise com os dados de crédito e de depósitos.

Se as predições do mercado de agências bancárias estiverem de acordo com a teoria, então nossa hipótese é que ele é geográfico. Analogamente, se no mercado de depósito ou de crédito as predições se confirmam, então existem evidências de que estes mercados abarcam tanto a dimensão do produto quanto geográfica.

⁵⁰ Vide por exemplo Nakane (2001) e Belaisch (2003).

3.3 BASE DE DADOS E RESULTADOS EMPÍRICOS

3.3.1 Base de dados

A fonte de dados principal deste trabalho é ESTBAN, um banco de dados do Banco Central do Brasil que armazena os saldos das principais contas contábeis por agência bancária de dez/2002. Utilizamos também os dados demográficos (população, área) do Atlas do Desenvolvimento Humano no Brasil do PNUD para o ano de 2000 e os dados do IBGE referentes às populações distribuídas nos bairros e subdistritos no mesmo ano.

No ESTBAN, o número total de bancos e de agências em dez/2002 era 118⁵¹ e 16.958, respectivamente. A variável depósito utilizada neste trabalho refere-se à soma dos depósitos de governo, privado, poupança e a prazo e a variável crédito, às operações de crédito realizada pelas agências. Na subseção seguinte utilizamos os dados desta base para esboçar as curvas de Lorenz.

Para derivar conclusões gerais neste estudo, o ideal é que o mercado esteja em equilíbrio. Pelos dados utilizados, nosso pressuposto é de que em dez/2002 o mercado atingiu o equilíbrio, o que parece ser razoável. O mercado bancário sofreu diversas reestruturações importantes antes desta data. Os anos 90 foram marcados pela estabilização da inflação, a partir da implementação do Plano Real em 1994, e também por um amplo processo de abertura da economia. Entre 1995-96, com o fim da inflação, o setor enfrentou diversos problemas e sua reestruturação contou com programas governamentais (PROER e PROES⁵²). O ingresso de bancos estrangeiros também foi significativo no final dos anos 90, através de compras de bancos de relativa importância. Entre eles podemos citar o HSBC com a aquisição

⁵¹ Alteramos os CPNJs de alguns bancos, considerando as principais fusões e aquisições ocorridas em período anterior a dez/2002, de modo a agrupar os bancos pertencentes aos principais conglomerados bancários. Desta forma, as seguintes mudanças foram feitas:

- 1) Alvorada, Mercantil de São Paulo, BCN, BEM e BEC => Bradesco;
- 2) Banerj, Banestado, BEG e BEMGE => Itaú;
- 3) BANDEPE e Sudameris => ABN AMRO Real; e
- 4) Santander Brasil, Banespa e Meridional => Santander.

⁵² PROER: Programa de Estímulo à Reestruturação e ao Fortalecimento do Sistema Financeiro Nacional – no âmbito do PROER o governo concedeu incentivos para a fusão, aquisição e reestruturação de instituições financeiras em dificuldades. PROES: Programa de Incentivos à redução do Setor Público Estadual na Atividade Bancária – em linhas gerais, foi um incentivo à reestruturação e às privatizações dos bancos públicos.

do Bamerindus (1997), o ABN AMRO que comprou o Banco Real (1998) e o Banco Santander que adquiriu o Banespa (2000). Os dados da **Tabela 3.1** da **Seção 3.1**, p.113 também confirmam a nossa hipótese, pois ali verificamos que a partir de 2002 o número de bancos não tem sofrido variações significativas.

A parte empírica é apresentada na seguinte seqüência: na **subseção 3.3.2**, esboçamos as curvas de Lorenz para diversas agregações territoriais; na **subseção 3.3.3** tentamos inferir qual é a divisão territorial que melhor representaria um submercado; uma forma de identificar submercados independentes em municípios com mais de um submercado é proposta na **subseção 3.3.4**; a visualização através de mapas das diversas agregações territoriais em submercados é feita na **subseção 3.3.5**; e por fim, na **subseção 3.3.6**, verificamos a distribuição do tamanho das firmas sob outra ótica, dividindo os bancos segundo a atuação no mercado em nacional e local.

3.3.2 Curvas de *Lorenz*

Nesta subseção vamos mostrar o formato das curvas de Lorenz⁵³ para o mercado bancário brasileiro. Nos gráficos, o eixo das abscissas representa a fração de bancos que atuam no mercado e o eixo das ordenadas, a razão de concentração. A coordenada (x, y) significa que a fração x de bancos possui participação de mercado de y . De acordo com a teoria, se o mercado em questão é composto por diversos submercados então a curva de Lorenz deve passar acima da curva-limite (**Equação 3.1**, p.121). Nos mercados altamente concentrados, as curvas de Lorenz são bastante assimétricas. Se um submercado é totalmente fragmentado, ou seja, todas as firmas exibem o mesmo tamanho, ele é representado por uma linha reta de 45° partindo da origem.

Para construir as curvas de Lorenz, utilizamos o volume de depósito, o volume de crédito e o número de agências bancárias por instituição financeira. Para cada um desses indicadores, calculamos a participação do mercado⁵⁴ de cada banco e as ordenamos de forma decrescente.

⁵³ As representações das curvas de Lorenz se referem à imagem espelhada da curva de Lorenz.

⁵⁴ Para cada mercado a participação de cada banco foi calculada da seguinte forma: depósitos e crédito=volume do banco dividido pelo volume total do mercado; e agências=número de agências do banco dividido pelo total de agências.

As curvas representam o *market share* que pertence à fração das maiores firmas que operam no mercado. Um banco que é o maior em volume de crédito em um mercado não necessariamente pode tê-lo em depósito ou agências. Esse procedimento foi repetido para cada nível de agregação (Brasil, Unidades da Federação, mesorregiões e alguns municípios)⁵⁵. Nos gráficos, a curva-limite é representada pela linha pontilhada preta e a reta cheia preta (quando houver) o mercado totalmente fragmentado. As curvas de Lorenz de depósito, crédito e agências têm cores verde, azul e vermelho, respectivamente.

Fonte dos dados: Banco Central do Brasil.

Inicialmente, analisamos a distribuição do tamanho das firmas considerando todo o território nacional. Na **Figura 3.2**, observa-se um mercado altamente concentrado onde 10% dos bancos possuem aproximadamente 91% do total de depósitos, 82% do total de empréstimo e 94% das agências bancárias. Se a indústria bancária brasileira é composta por vários submercados, então, de acordo com o modelo, as curvas de Lorenz para cada indicador devem satisfazer o limite inferior determinado pela **Equação 3.1**. Na **Figura 3.2**, verifica-se que essas curvas passam bem acima da curva-limite.

Na **Figura 3.3** apresentamos as curvas de Lorenz das 27 unidades da Federação para o volume de depósito, volume de crédito e número de agências, separadamente. Para cada gráfico, exibimos também as respectivas curvas do Brasil da **Figura 3.2** (as curvas mais grossas, e em geral, as mais assimétricas). Para os três indicadores, observa-se que para a

⁵⁵ De acordo com o critério de divisão geográfica adotada pelo IBGE (2000), o Brasil está dividido em 26 Estados mais o Distrito Federal / 137 mesorregiões / 558 microrregiões / 5.507 municípios.

maioria das UFs as curvas de Lorenz passam entre a curva-limite e a curva de Lorenz para a agregação do Brasil. Neste nível de desagregação, as firmas em operação também exibem tamanhos distintos e se as previsões do modelo de Sutton são válidas, essa desigualdade advém da junção de diversos submercados independentes.

Figura 3.3 – Curva de Lorenz para UF - depósito (esquerda superior), crédito (direita superior) e agências (esquerda inferior)

A linha contínua mais grossa refere-se à curva do Brasil (**Figura 3.2**). A linha pontilhada é a curva-limite.

Fonte dos dados: Banco Central do Brasil.

É provável que quanto maior a desagregação territorial, maior a probabilidade de emergir um submercado. Por isso, seguimos com a investigação utilizando a divisão por mesorregiões adotada pelo IBGE. Para verificar a concentração do mercado, a título de ilustração, apresentamos na **Figura 3.4** as curvas de Lorenz das mesorregiões do estado de Pernambuco bem como para todo o estado (para outras UFs vide **APÊNDICE 3.A**, p.155).

Figura 3.4 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Mesorregiões de Pernambuco

As linhas mais grossas referem-se às respectivas curvas de Lorenz para o estado de Pernambuco. A linha pontilhada em preto é a curva-limite. As linhas pontilhadas mais claras são as curvas de Lorenz para as mesorregiões.

Fonte dos dados: Banco Central do Brasil.

Para este nível de desagregação, observa-se que para a maioria dos casos (vide **Figura A - 3.1-Figura A - 3.3** do **APÊNDICE 3.A**, p.155), as curvas de Lorenz estão limitadas superiormente pela respectiva curva da UF (mantendo a mesma relação encontrada na **Figura 3.3**). O limite inferior, por sua vez, é violado mais freqüentemente (principalmente para agências bancárias) havendo casos em que o mercado se apresenta quase totalmente fragmentado. Essa variação na localização das curvas (acima ou abaixo da curva-limite) pode ser decorrente das diferenças sócio-econômicas, demográficas e geográficas que, por sua vez, determinam o número de submercados. Regiões populosas devem concentrar vários submercados enquanto regiões desabitadas devem constituir um único submercado. Esse fato é evidenciado quando efetuamos a desagregação por tamanho dos municípios. Na **Figura 3.5** mostramos as curvas para os municípios com até 10 mil habitantes e pelo menos dois bancos (32 municípios). Na maioria deles, as curvas de Lorenz de depósito e de crédito estão abaixo da curva-limite e para 100% dos casos, o mercado de agências bancárias é totalmente fragmentado.

Figura 3.5 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Municípios com dois ou mais bancos e até 10 mil habitantes

Fonte dos dados: Banco Central do Brasil.

Quando selecionamos os municípios com mais de 500 mil habitantes (31 municípios), os mercados apresentam-se mais concentrados e a maioria das curvas passa acima da curva-limite (vide **Figura 3.6**):

Figura 3.6 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Municípios com mais de 500 mil habitantes

Fonte dos dados: Banco Central do Brasil.

Os resultados iniciais sugerem que o mercado bancário considerando o Brasil como um todo ou as UFs é composto de vários submercados, exibindo grande desigualdade no tamanho dos bancos. À medida que fazemos as desagregações territoriais, as firmas apresentam tamanhos mais simétricos, sendo que este resultado é mais forte para as agências do que para depósito e crédito.

No entanto, ainda não está claro qual é o limite geográfico de um submercado. O Brasil apresenta diferenças geográficas muito marcantes e por isso, para alguns lugares as mesorregiões poderiam ser consideradas um submercado, enquanto para outros poderiam ser

as microrregiões ou os municípios. Na subseção seguinte tentaremos encontrar evidências que nos permitam identificar qual é a extensão geográfica que melhor representaria um submercado.

3.3.3 Identificando o tamanho dos submercados

O objetivo desta subseção é investigar qual é a extensão geográfica mais apropriada para a definição de submercado, apresentando alguns resultados empíricos. Conforme visto na **subseção 3.3.2 (Figura 3.5 e Figura 3.6, p.131)**, cidades pequenas parecem representar um submercado, enquanto cidades grandes parecem comportar diversos deles. Escolhemos as microrregiões e os municípios para identificar se eles representam ou não um submercado.

O número de bancos (*nbanks*) que opera em cada mercado deve ser função do número de agências, da intensidade de competição e dos custos de entrada. Em particular, nossa hipótese é que se uma microrregião (município) retrata um submercado, então o número de bancos deve ser influenciado positivamente pelo número de agências que aí atuam (*nbranch*) e negativamente (ou não influenciadas, ou ainda com influência comparativamente menor a *nbranch*) pelo número de agências que atuam em mercados adjacentes (*adbranch*).

A variável dependente assume valores ordenados e discretos $\{0,1,2,...,J\}$ e pode ser modelada como um *logit* ordenado (vide Wooldridge, 2002, p. 504). $nbanks^*$ é uma variável latente determinada por:

$$\text{Equação 3.2: } nbanks^* = \beta_1 nbranch_s + \beta_2 adbranch_s + X'_s \beta_3 + \varepsilon_s = X' \beta + \varepsilon_s$$

Onde $s=1,...,S$ é o submercado (microrregião ou município), X' é o vetor de variáveis independentes, β é o vetor de coeficientes a serem estimados e ε_s é o erro não observado tal que ε_s / X tem distribuição logística com média zero e variância $\pi^2/3$. Seja $\alpha_1 < \alpha_2 < ... < \alpha_J$ os pontos de corte (*cut points*) e defina:

$$nbanks = 0 \text{ se } nbanks^* \leq \alpha_1$$

$$nbanks = 1 \text{ se } \alpha_1 < nbanks^* \leq \alpha_2$$

...

$$nbanks = J \text{ se } nbanks^* > \alpha_J$$

Seja $\Lambda(.) \equiv \frac{\exp(.)}{1 + \exp(.)}$ a função *logit*. Para cada observação i , a função de verossimilhança é escrita como:

$$l_i(\alpha, \beta) = 1[nbanks_i = 0] \log[\Lambda(\alpha_1 - X_i\beta)] + 1[nbanks_i = 1] \log[\Lambda(\alpha_2 - X_i\beta) - \Lambda(\alpha_1 - X_i\beta)] + \dots + 1[nbanks_i = J] \log[1 - \Lambda(\alpha_J - X_i\beta)]$$

E os parâmetros α e β podem ser estimados por máxima verossimilhança.

Como $nbanks^*$ é uma variável latente, uma possível interpretação dos coeficientes é feita através da padronização dos seus desvios-padrão estimados. O $nbanks^*$ -standardized coefficient para X_k é dado por $\beta_k^{nbanks^*} = \frac{\beta_k}{\sigma_{nbanks^*}}$ e o fully standardized coefficient é

$$\beta_k^S = \frac{\sigma_k \beta_k}{\sigma_{nbanks^*}}, \text{ onde } k \text{ é o índice da variável independente, } \sigma_{nbanks^*} \text{ é o desvio-padrão}$$

estimado de $nbanks^*$ e σ_k o desvio-padrão estimado da variável independente k .

Para mitigar os efeitos de uma microrregião (município) grande e que provavelmente comporta diversos submercados, procedemos à sua eliminação, juntamente com as microrregiões (municípios) localizadas nas suas proximidades. No caso de microrregiões, inicialmente retiramos da amostra as que possuíam mais de 1 milhão de habitantes. Após, excluimos aquelas que estivessem a uma distância⁵⁶ (entre os centros geográficos⁵⁷) inferior a

⁵⁶ A distância entre duas coordenadas geográficas foi calculada pela fórmula do *Great-circle distance*. Seja $(\phi_s, \lambda_s); (\phi_f, \lambda_f)$ as coordenadas geográficas da latitude e longitude de dois pontos em radianos. A distância do grande círculo foi calculada como: $\Delta dist = 6.372,795x \arcsin \left(\sqrt{\sin^2 \left(\frac{\phi_f - \phi_s}{2} \right) + \cos \phi_s \cos \phi_f \sin^2 \left(\frac{\lambda_f - \lambda_s}{2} \right)} \right)$.

75 km dessas microrregiões com mais de 1 milhão de habitantes. Analogamente, eliminamos os municípios que possuíam mais de 350 mil habitantes e aqueles cuja distância (entre as sedes) fosse inferior a 50 km desses municípios com mais de 350 mil habitantes. Fazendo essas exclusões, na primeira amostra restaram 487 microrregiões e na segunda, 4.847 municípios.

Para cada microrregião (município), obtivemos o número de bancos, bem como o número de agências bancárias. Controlamos pelo tamanho do mercado através das variáveis área e população. A concentração do mercado é capturada pelo índice de *Herfindahl* do volume de depósitos ou empréstimos ($hdep, hloan$). A intensidade de competição é representada pelo número de agências dividido pela população. A razão entre o desvio-padrão e a média do volume de empréstimos captura o grau de diferenciação presente no mercado. Adicionalmente, utilizamos *dummies* para as regiões geográficas (NE, CO, SE, S).

Para cada microrregião, computamos adicionalmente o número de agências bancárias e a população das demais microrregiões cuja distância fosse inferior ou igual a 100 km. No caso de municípios, para aqueles cuja distância fosse menor ou igual a 60 km e 30 km, respectivamente para as regiões norte/centro-oeste e nordeste/sul/sudeste. Esses valores compõem as variáveis *adbranch* e *ladpopt*, respectivamente.

A fonte dos dados são o ESTBAN do Bacen (número de agências e volume de empréstimos em dez/2002), o Atlas do Desenvolvimento Humano no Brasil do PNUD (para os dados populacionais e demográficos do ano 2000). A descrição das variáveis utilizadas está na **Tabela 3.4:**

⁵⁷ Para os municípios, as coordenadas da latitude e da longitude referem-se à sede do município. Para as microrregiões, calculamos o centro da coordenada geográfica como a média das latitudes e longitudes das sedes dos municípios pertencentes à microrregião. Neste caso, o “centro” pode não necessariamente estar localizado dentro da microrregião. Quando afirmamos que duas microrregiões (municípios) distam 100 km uma da outra, significa que a distância entre os seus respectivos centros geográficos (sedes) é de 100 km.

Tabela 3.4 – Descrição das variáveis

Variável	Descrição
<i>nbranch</i>	Número de agências bancárias na microrregião (município).
<i>adbranch</i>	Número de agências bancárias nas microrregiões (municípios) adjacentes. Para dada microrregião, uma microrregião adjacente é aquela cuja distância entre os centros é igual ou inferior a 100 km. Computa-se então o número de agências bancárias de todas essas microrregiões adjacentes. Para os municípios, o cálculo é análogo, diferindo apenas na distância. Um município é adjacente a outro se a distância é menor ou igual a 60 km (NO e CO) ou 30 km (NE, SE, S).
<i>lpopt</i>	$\ln$ (logaritmo natural) da população residente total (mil habitantes) da microrregião (município).
<i>ladpopt</i>	$\ln$ da população residente total (mil habitantes) da microrregião (município) adjacente. A população adjacente foi definida como 0,1 quando igual a zero, para que o $\ln$ fosse definido nesta situação. O cálculo dessa variável é análogo a <i>adbranch</i> .
<i>dens</i>	Densidade populacional da microrregião (município) (=população/área).
<i>hdep, hloan</i>	<i>hdep</i> : Índice <i>Herfindahl</i> do volume de depósitos calculado para a microrregião (município), definido como: $hdep = \sum_{i=1}^n s_i^2$, onde s_i^2 é o quadrado da proporção do volume de depósitos do banco i , na microrregião (município). <i>hloan</i> : idem para o volume de empréstimos.
<i>area</i>	Área geográfica da microrregião (município) em Km ² /1000.
<i>nagencph</i>	Número de agências por mil habitantes na microrregião (município).
<i>loancv</i>	Coeficiente de variação do volume de empréstimos no município (=desvio-padrão/média).
<i>dummies</i>	Variáveis <i>dummies</i> (=1) para a região (NE, CO, SE, S) a que pertence a microrregião (município). A região NO é a categoria-base.

Na nossa amostra, poucas microrregiões apresentam nenhum banco (2,7% ou 13 de 487). Para os municípios, esse percentual é bastante elevado, correspondendo a 43,1% (2.090 de 4.847) da amostra, conforme observado na **Tabela 3.5**:

Tabela 3.5 – Número de bancos

<i>nbanks</i>		0	1	2	3	4	5	6	7	8	≥9	total
Nº	Microrregião	13	25	61	58	81	64	41	31	38	75	487
obs.	Município	2.090	1.270	568	279	233	142	75	65	48	77	4.847

Fonte dos dados: Bacen (dez/2002).

Nos mercados que não possuem bancos, as variáveis *nbranch* e *nagencph* são iguais a zero e o Índice de *Herfindahl* não está definido. Por isso, optamos por excluí-los de nossa amostra.

A estatística descritiva das variáveis utilizadas ($nbanks^* > 0$) se encontra na **Tabela 3.6**:

Tabela 3.6 – Estatística descritiva das variáveis

Variável	Microrregião				Município			
	Média	Desv. Padr.	Min.	Máx.	Média	Desv. Padr.	Min.	Máx.
<i>nbanks</i>	5,08	2,50	1,00	9,00	2,49	2,04	1,00	9,00
<i>nbranch</i>	17,79	19,31	1,00	155,00	2,86	3,93	1,00	73,00
<i>adbranch</i>	74,20	87,42	0,00	414,00	11,79	14,00	0,00	125,00
<i>popt</i>	198,91	160,73	2,05	977,19	26,45	38,13	0,80	345,97
<i>adpop</i>	746,20	721,22	0,10	3.990,28	99,38	106,43	0,10	1.093,70
<i>dens</i>	42,85	109,56	0,22	1.983,74	48,85	111,87	0,20	3.284,32
<i>hdep</i>	0,40	0,21	0,14	1,00	0,70	0,31	0,13	1,00
<i>area</i>	16,75	31,51	0,02	333,86	1,90	6,95	0,00	161,45
<i>nagencph</i>	0,09	0,06	0,01	0,49	0,15	0,10	0,02	1,26
<i>loancv</i>	-	-	-	-	0,94	0,36	0,02	4,83

Fonte dos dados: Bacen (dez/2002) e Pnud (2000). N° obs.: 474 microrregiões e 2.757 municípios; 1.487 *loancv*.

Em média uma microrregião da amostra apresenta 5 bancos, 18 agências, população de 199 mil habitantes e em suas adjacências existem em média 74 agências e 745 mil habitantes. Um município médio da amostra, por sua vez, apresenta 2 bancos, 3 agências, 12 agências nos municípios adjacentes e população no município e nas proximidades de 26 mil e 99 mil habitantes, respectivamente. A concentração (*hdep*) e a intensidade de competição (*nagencph*) são, em média menores nas microrregiões do que nos municípios.

Para diminuir os problemas na estimação, verificamos quais variáveis independentes apresentavam correlação alta dentro de uma categoria da variável dependente. Desta forma, excluímos a variável *nagencph* da amostra de municípios.

As estimativas do *logit* ordenado são apresentadas na **Tabela 3.7**. A coluna (1) refere-se às microrregiões e as demais aos municípios. Em (2), consideramos a amostra total e em (3) excluímos os municípios com apenas um banco, devido à grande proporção na amostra⁵⁸. Na parte superior aparecem os coeficientes estimados e na parte inferior os valores dos pontos de corte (*cut points*). O *Wald chi2* testa se todos os coeficientes com exceção dos pontos de corte são iguais a zero. Em todos os casos, os p-valores (0,00) indicam que rejeitamos a hipótese nula a 1% de significância.

⁵⁸ A teoria apresentada deriva predições para o agregado de submercados, podendo em particular ser composto por diversos submercados que comportam uma única firma. No caso em que o submercado é composto por uma firma, ao mesmo tempo em que ele é totalmente fragmentado, é altamente concentrado. Por sua vez, os municípios que não possuem banco podem ser um submercado ou fazer parte de outro. No presente caso, estimamos considerando e não considerando os municípios com apenas um banco. Para as microrregiões, a proporção desses casos é bem menor, não devendo causar maiores problemas.

Tabela 3.7 – Resultados das estimativas do modelo *logit* ordenado

Variável	(1) Microrregiões		(2) Municípios		(3) Municípios	
	Coef.	Desv. Padr.	Coef.	Desv. Padr.	Coef.	Desv. Padr.
<i>nbranch</i>	0,0655	0,0218 *	2,9965	0,3994 *	2,7262	0,3785 *
<i>adbranch</i>	-0,0011	0,0020	-0,0005	0,0073	-0,0015	0,0073
<i>lpopt</i>	2,0854	0,3672 *	0,6662	0,1972 *	0,5150	0,2150 &
<i>ladpopt</i>	0,0226	0,0447	-0,0281	0,0502	-0,0264	0,0519
<i>dens</i>	0,0021	0,0022	-0,0013	0,0022	-0,0010	0,0022
<i>hdep</i>	-23,1530	2,0306 *	-28,5145	2,6318 *	-26,9496	2,7257 *
<i>area</i>	-0,0052	0,0045	-0,0151	0,0066 &	-0,0156	0,0062 &
<i>nagencph</i>	4,6621	4,1324	-	-	-	-
<i>loancv</i>	-	-	-	-	1,5524	0,4461 *
<i>/cut1</i>	-12,5459	2,3784	-18,1059	2,7674	-2,3623	1,7508
<i>/cut2</i>	-3,7190	1,9296	-3,4261	1,8551	4,0479	1,7313
<i>/cut3</i>	-0,3088	1,9349	3,6247	1,8562	9,1825	1,9331
<i>/cut4</i>	2,7010	1,9661	9,1841	2,0816	13,6299	2,2203
<i>/cut5</i>	4,5913	1,9808	13,9545	2,4000	17,7461	2,5748
<i>/cut6</i>	6,0401	1,9857	18,3903	2,7859	22,7020	2,9837
<i>/cut7</i>	7,2803	1,9870	23,9450	3,2549	28,1888	3,5176
<i>/cut8</i>	8,8617	1,9799	30,0876	3,8510	-	-
<i>Dummies</i>	NE, CO, SE, S		NE, CO, SE, S		NE, CO, SE, S	
<i>Nº obs.:</i>	474		2.757		1.487	
<i>Wald chi2:</i>	397,06 (0,00)		464,92 (0,00)		361,48 (0,00)	
<i>LR Test</i>	81,99 (0,54)		763,80 (0,00)		692,40 (0,00)	

A variável dependente *nbanks*^{*}>9 foi substituída por 9. Desvio-padrão robusto à heterocedasticidade.

P-valor do *Wald chi2* e *LR test* entre parênteses. Coeficientes das *dummies* não reportado.

*, & e # indicam coeficiente da variável associada significativo a 1%, 5% e 10%, respectivamente.

Nas três estimações (**Tabela 3.7**), os coeficientes do número de agências (*nbranch*) no submercado são positivos e significativos e os coeficientes das agências adjacentes (*adbranch*) são negativos e não-significativos. Os coeficientes da população (*lpopt*) são positivos e significativos e para a população adjacente (*ladpopt*) são não-significativos. A interpretação dos sinais dos coeficientes é completada com o cálculo das variações marginais da variável dependente apresentadas na **Tabela 3.8**:

Tabela 3.8 – Variação marginal de *nbanks**

Variável	(1) Microrregiões		(2) Municípios		(3) Municípios	
	$\beta_k^{Snbank*}$	β_k^S	$\beta_k^{Snbank*}$	β_k^S	$\beta_k^{Snbank*}$	β_k^S
<i>nbranch</i>	0,0091	0,1749	0,1577	0,6196	0,1643	0,7910
<i>adbranch</i>	-0,0001	-0,0128	0,0000	-0,0004	-0,0001	-0,0015
<i>lpopt</i>	0,2885	0,2199	0,0351	0,0343	0,0310	0,0269
<i>ladpopt</i>	0,0031	0,0122	-0,0015	-0,0027	-0,0016	-0,0029
<i>dens</i>	0,0003	0,0313	-0,0001	-0,0074	-0,0001	-0,0091
<i>hdep</i>	-3,2033	-0,6576	-1,5005	-0,4667	-1,6238	-0,3051
<i>area</i>	-0,0007	-0,0228	-0,0008	-0,0055	-0,0009	-0,0074
<i>nagencph</i>	0,6450	0,0393	-	-	-	-
<i>loancv</i>	-	-	-	-	0,0935	0,0336
<i>Dummies</i>	NE, CO, SE, S		NE, CO, SE, S		NE, CO, SE, S	
<i>Nº obs.:</i>	474		2.757		1.487	

A variável dependente *nbanks** > 9 foi substituída por 9. Coeficientes das *dummies* não reportado.

$\beta_k^{Snbank*}$: *nbanks** -standardized coefficient. β_k^S : fully standadized coefficient.

Na coluna (1) da **Tabela 3.8**, o valor $\beta_{nbranch}^{Snbanks*}$ é interpretado da seguinte maneira: tudo o mais constante, aumentando-se uma agência no mercado, espera-se que *nbanks** amplie em 0,0091 desvio-padrão. Analogamente, $\beta_{nbranch}^S$ indica que mantendo as demais variáveis constante, aumentando-se um desvio-padrão na variável *nbranch*, espera-se que o número de bancos no mercado cresça em 0,1749 desvio-padrão. Para os demais coeficientes, a interpretação segue do mesmo modo.

Para as três estimações, os coeficientes mostram o seguinte padrão (**Tabela 3.8**): o número de bancos que atuam na microrregião (município) é fortemente influenciado pelas agências existentes e pela sua população e não influenciados pelas agências e população de localidades vizinhas. De acordo com a nossa hipótese, os resultados sugerem que as microrregiões (municípios), excluídos os mais populosos e as respectivas localidades adjacentes, podem ser considerados submercados independentes.

Uma hipótese implícita dos modelos *logit* ordenado é a da verossimilhança proporcional (*proportional odds assumption, parallell regression*) (vide Long e Freese, 2006, p. 197). Este modelo é equivalente a $J - 1$ regressões binárias com o pressuposto de que as inclinações dos coeficientes são idênticas. Para verificar se isso ocorre nos dados utilizados computamos o *LR test* através do comando **omodel**⁵⁹ disponível para o Stata. Na última linha da **Tabela 3.7**, o p-

⁵⁹ Este comando não é parte oficial do Stata e foi desenvolvido por Wolfe e Gould (1998), vide Long e Freese (2006), p.199. O teste é calculado sem uso da variância robusta à heterocedasticidade

valor de 0,54 da estimação (1) indica que não violamos a hipótese de verossimilhança proporcional. No entanto, em (2) e em (3), o p-valor de 0,00 traz evidência de que a hipótese de regressão paralela pode ser rejeitada a 1% de significância. Para os municípios este resultado revela que as relações entre as diversas categorias não são proporcionais, sendo necessário outro modelo para a estimação.

O *generalized ordered logit model* permite que os coeficientes sejam diferentes para cada uma das $J - 1$ regressões e a sua implementação pode ser feita pelo comando **gologit2**⁶⁰ disponível para o Stata. A vantagem dessa metodologia com relação ao *logit* ordenado é o fato de ser menos restritiva e a desvantagem é que inclui-se muito mais parâmetros a serem estimados. Para diminuir o problema de existência de categorias da variável dependente com poucas observações, redefinimos $nbanks^* = 6$, se $nbanks^* = 7$ e $nbanks^* = 7$, se $nbanks^* \geq 8$. A variável *dummy* foi definida igual a 1 para as regiões SE e S e zero caso contrário. As estimações apresentadas na **Tabela A - 3.1** e **Tabela A - 3.2** do APÊNDICE 3.B, p.158 equivalem a uma série de regressões logísticas binárias. Na **Tabela A - 3.1**, a primeira coluna refere-se a estimação da categoria 1 versus as categorias 2-7; na segunda coluna, às categorias 1-2 versus 3-7 e assim por diante. Os coeficientes positivos da variável *nbranch* indicam que quanto maior o seu valor, maior a probabilidade de estar em uma categoria mais alta. Os coeficientes negativos de *lpopt* indicam que valores maiores desta variável aumentam a verossimilhança de pertencer à categoria mais baixa. Por sua vez, *adbranch* é quase sempre não-significativo. Na implementação, deparamos freqüentemente com probabilidades preditas negativas⁶¹, o que faz com essa técnica seja menos atraente que o *logit* ordenado.

3.3.3.1 Identificando submercados na região CO

Outra forma utilizada para identificar os submercados foi separando-os em diversas regiões geográficas (mesorregião, microrregião, município e bairros), conforme o caso. Realizamos esta divisão para os estados da região centro-oeste, exceto o Distrito Federal. O critério utilizado para incluir uma região como submercado é que viole o limite inferior de Sutton

⁶⁰ Este comando foi desenvolvido por Richard Williams, vide <http://www.nd.edu/~rwilliam/gologit2/>.

⁶¹ Apesar de retirarmos variáveis e a adicionarmos gradualmente, ainda assim não conseguimos mitigar este problema. A predição de probabilidade negativa pode ocorrer quando as linhas se cruzam, vide <http://www.nd.edu/~rwilliam/gologit2/>.

para ao menos dois critérios⁶². Inicialmente, escolhemos as mesorregiões, depois as microrregiões e por fim os municípios que não atendem ao limite.

Entre os municípios do centro-oeste que atendem ao limite inferior de Sutton e para os quais possuímos dados detalhados da população por bairro estão Campo Grande, Cuiabá, Várzea Grande e Goiânia. Para eles, utilizamos a análise de cluster para separá-los em diversos submercados (vide maiores detalhes na **subseção 3.3.4**, p.142). Feito isto, para a região centro-oeste computamos 212 submercados, sendo 4 mesorregiões, 28 microrregiões, 156 municípios, 6 em Campo Grande, 8 em Cuiabá, 3 em Várzea Grande e 7 em Goiânia. Do total, 50 municípios não possuíam agências e por isso foram excluídos da amostra, restando 162 observações⁶³.

Os submercados adjacentes são aqueles cuja distância entre os centros⁶⁴ é de até 50 km. Para os submercados das capitais e Várzea Grande, a população foi calculada com os dados do IBGE para bairros e subdistritos baseado na localização da agência. A área total do município foi dividida em partes iguais para cada submercado.

Cerca de 1/3 da amostra possui apenas um banco. Em média, cada submercado possui 3 bancos, 6 agências e 20 agências na vizinhança. A estatística descritiva das variáveis bem como o número de bancos por submercado se encontram na **Tabela 3.9**:

⁶² Incluímos também algumas microrregiões que possuíam vários municípios com 0 e 1 bancos.

⁶³ Nessa configuração de agrupamento existem diversos submercados com uma ou nenhuma agência. Para tentar contornar este problema, estimamos (resultado não-reportado) modelo similar agregando esses municípios dentro de uma microrregião, sempre que possível, de modo a diminuir os submercados com 0 ou 1 agência. A estimação do *logit* ordenado neste caso, traz resultados menos satisfatórios do que os apresentados nesta subseção.

⁶⁴ A distância entre os submercados foi calculada da mesma forma que anteriormente. Para os municípios, as coordenadas de latitude e a longitude referem-se à sede do município. Para as mesorregiões e microrregiões calculamos o centro da coordenada geográfica como a média das latitudes e longitudes das sedes dos municípios pertencentes à região. Para os bairros é a média das coordenadas das agências bancárias.

Tabela 3.9 – Estatística descritiva das variáveis e número de bancos - CO

Variável	Média	Desv. Padr.	Min.	Máx.	<i>nbanks</i>	n° obs.
<i>nbanks</i>	3,19	2,46	1,00	13,00		
<i>nbranch</i>	5,99	8,36	1,00	46,00	1	51
<i>adbranch</i>	19,80	35,70	0,00	176,00	2	33
<i>popt</i>	3,29	1,28	0,74	6,15	3	21
<i>adpop</i>	2,71	3,47	-2,30	7,20	4	16
<i>dens</i>	114,45	429,58	0,61	3.417,81	5	13
<i>hdep</i>	0,62	0,29	0,18	1,00	6	15
<i>hloan</i>	0,75	0,25	0,13	1,00	≥ 7	13
<i>area</i>	9,56	20,32	0,06	124,64	total	162

Fonte dos dados: Bacen (dez/2002), IBGE (2000) e Pnud (2000).

Os resultados da estimativa por *logit* ordenado estão em consonância com os apresentados para as microrregiões e para os municípios (vide **Tabela 3.10**). O coeficiente do número de agências no submercado é positivo e significativo e do número de agências adjacente ao submercado é não significativo. O número de agências em um submercado é influenciado positivamente pela sua população e negativamente pela população adjacente.

O valor $\beta_{nbranch}^{Snbanks^*}$ da **Tabela 3.10** indica que tudo o mais constante, aumentando-se uma agência no mercado, espera-se que $nbanks^*$ amplie em 0,016 desvio-padrão. E resultado análogo ao obtido anteriormente se mantém: o número de bancos que atuam no mercado é fortemente influenciado pelas agências existentes e pela sua população, não influenciado pelas agências adjacentes e negativamente influenciado pela população na vizinhança. Pelo p-valor do teste *LR*, não rejeitamos a hipótese de regressão paralela a 1% de significância.

Tabela 3.10 – Resultados das estimativas do modelo *logit* ordenado para os submercados da região CO

	Coefficiente		Desv. Padr.	$\beta_k^{Snbank^*}$	β_k^S
<i>nbranch</i>	0,0950	#	0,0513	0,0160	0,1341
<i>adbranch</i>	0,0083		0,0065	0,0014	0,0502
<i>lpopt</i>	1,1355	*	0,2323	0,1917	0,2453
<i>ladpopt</i>	-0,2728	*	0,0744	-0,0460	-0,1599
<i>dens</i>	-0,0007	#	0,0004	-0,0001	-0,0486
<i>hloan</i>	-16,0456	*	1,9942	-2,7087	-0,6805
<i>area</i>	-0,0268	&	0,0121	-0,0045	-0,0920
<i>Dummies</i>	MT, GO				
<i>Wald chi2:</i>	139,97 (0,00)				
<i>LR Test</i>	126,84 (0,01)				

A variável dependente é $nbanks^*$. N°obs.: 162. Desvio-padrão robusto à heterocedasticidade. P-valor do *Wald chi2* e *LR test* entre parênteses. Coeficientes das *dummies* não reportado.

*, & e # indicam coeficiente da variável significativo a 1%, 5% e 10%, respectivamente.

$\beta_k^{Snbank^*}$: $nbanks^*$ -standardized coefficient. β_k^S : fully standadized coefficient.

A partir das evidências empíricas obtidas nesta subseção, podemos tirar as seguintes conclusões: 1) as microrregiões geograficamente isoladas podem ser consideradas submercados de agências bancárias; 2) idem para os municípios. No entanto, para este caso, existem algumas ressalvas. O fato de existir muitos municípios pequenos (com uma ou nenhuma agência) cujos índices de concentração não estão definidos ou não apresentam variabilidade pode comprometer as estimativas que os utilizem como *proxy* de submercados. Além disso, rejeitou-se a hipótese de regressão paralela; e 3) a divisão de um estado em submercados que englobam mesorregiões, microrregiões, municípios e bairros apresenta resultados análogos aos da microrregião.

3.3.4 Identificando submercados independentes em municípios com mais de um submercado

Se considerarmos que os submercados podem ser identificados pela relativa proximidade geográfica das agências bancárias, então é possível diferenciar os vários existentes dentro de uma região geográfica. Nossa hipótese é que se duas agências estão localizadas próximas uma da outra, então elas pertencem a um mesmo submercado e, se estão longe, então pertencem a submercados distintos. A análise de cluster pode ser utilizada para identificar os submercados numa cidade com diversos deles.

As variáveis utilizadas para a formação dos clusters são as coordenadas geográficas⁶⁵ (latitude - *lat* e longitude - *lon*) de cada agência bancária no município. Os métodos utilizados foram o *complete linkage* e o *average linkage*. Em ambos os casos, a proximidade entre os clusters é calculada pela distância Euclidiana. No *complete linkage* é calculada a distância entre o par de observações mais remoto, onde um dos pares está no primeiro cluster e o outro par, no segundo. No *average linkage* é calculada a distância média entre todos os pares de observações onde um dos pares está no primeiro cluster e o outro está no segundo (Rabe-Hesketh e Everitt, 2004).

⁶⁵ As coordenadas foram extraídas do *Google Maps* (<http://maps.google.com/>) e foram aproximadas quando a localização exata não foi encontrada.

Em particular, trabalhamos com os municípios do centro-oeste, cujo resultado foi utilizado na subseção anterior (Campo Grande, Cuiabá, Várzea Grande e Goiânia). A **Tabela 3.11** traz alguns dados acerca dessas cidades:

Tabela 3.11 – Dados dos municípios selecionados

Município	n° de bancos	n° de agências	população
Campo Grande	14	75	663.621
Cuiabá	13	56	483.346
Várzea Grande	6	11	215.298
Goiânia	14	153	1.093.007

Fonte dos dados: Bacen (dez/2002) e Pnud (2000).

A seguir detalharemos o processo para a obtenção dos clusters para Campo Grande, sendo que os demais municípios seguiram o mesmo procedimento. Inicialmente, obtemos as coordenadas geográficas de cada agência bancária da cidade e, na sequência, obtemos o dendograma:

Figura 3.7 – Dendograma para a análise de cluster de Campo Grande – *complete linkage* (esquerda) e *average linkage* (direita)

Na **Figura 3.7** à esquerda, apresentamos o dendograma do *complete linkage* e à direita, do *average linkage*. Ambas figuras nos mostram evidência da estrutura de cluster nos dados, apesar de ela ser um pouco diferente entre os dois métodos. A estrutura aninhada sugere a existência de diferentes formas de agrupar as agências e a questão é determinar o número de grupos que melhor se ajusta aos dados. Utilizamos o Índice de Duda e Hart como critério para encontrar os clusters existentes (vide **Tabela 3.12**).

Tabela 3.12 – Índice de Duda e Hart

Número de clusters	Duda/Hart			
	<i>Complete linkage</i>		<i>Average linkage</i>	
	<i>Je(2)/Je(1)</i>	<i>Pseudo T2</i>	<i>Je(2)/Je(1)</i>	<i>Pseudo T2</i>
1	0,8625	11,64	0,8625	11,64
2	0,6761	34,49	0,6180	44,50
3	0,2543	64,50	0,5689	46,23
4	0,3276	98,53	0,3348	93,36
5	0,2944	26,36	0,2422	28,15
6	0,3005	20,95	0,7635	3,41

Geralmente clusters distintos podem ser identificados pelo grande valor do Índice Duda e Hart e pequeno valor do *pseudo T2*. Adotando esse critério, o resultado do *complete linkage* não sugere nenhuma estrutura de cluster, enquanto o *average linkage* aponta para 6 grupos. Para os municípios de Cuiabá, Várzea Grande e Goiânia encontramos 8, 3 e 7 clusters, respectivamente. De acordo com a nossa hipótese, cada cluster pode ser considerado um submercado. Na **Figura 3.8**, mostramos a dispersão das agências conforme a latitude e a longitude e a identificação dos seis clusters para Campo Grande.

Figura 3.8 – Matriz com as observações identificadas pelo número do cluster (*average linkage*) – Campo Grande - MS

A localização aproximada desses seis clusters da cidade de Campo Grande pode ser visualizada no mapa da cidade na **Figura A - 3.4** do **APÊNDICE 3.C**, p.160. Para os quatro

municípios, a população de cada submercado utilizada na **subseção 3.3.3.1** foi baseada nos bairros localizados na sua proximidade, com o “resto” atribuído ao submercado da zona central.

Na **Figura 3.9** apresentamos as curvas de Lorenz de depósito, crédito e agências bancárias para os seis clusters identificados no município de Campo Grande. A linha contínua representa a curva de Lorenz do município, a linha tracejada em preto a curva-limite e as linhas tracejadas coloridas as curvas de Lorenz para cada um dos submercados. Verifica-se que para as agências bancárias, os seis submercados não satisfazem o limite inferior. Para depósito e crédito, por sua vez, aproximadamente três submercados se enquadram nesta categoria.

Figura 3.9 – Curva de Lorenz de depósito (esquerda), crédito (centro) e agências bancárias (direita) – Campo Grande - MS

Fonte dos dados: Banco Central do Brasil.

Em geral, as curvas de Lorenz para as agências bancárias tendem a violar o limite inferior de Sutton mais frequentemente do que para os depósitos ou crédito. Isso pode ser um indício de que o mercado de agências bancárias é geográfico e o mesmo não pode ser afirmado para depósito ou crédito⁶⁶. De outra maneira, a desagregação de um município em diversos submercados não necessariamente implica que encontrar-se-ão submercados de depósito ou crédito (vide o exemplo de municípios com até 10 mil habitantes, onde em alguns casos, a curva de Lorenz para crédito passa acima da curva-limite - **Figura 3.5**, p.131). Esse fato está em consonância com o discutido na **Seção 3.2.4**, p.124): é menos provável que o mercado de depósitos e crédito seja local. Se os bancos veiculam campanhas estaduais ou regionais que

⁶⁶ Curvas com características similares às da **Figura 3.9** foram obtidas para os municípios de Cuiabá, Goiânia e Porto Velho.

estimulam as pessoas a tomarem empréstimos, é mais difícil sustentar a existência de submercados que sejam independentes.

3.3.5 Mapas dos submercados de agências bancárias

O objetivo desta subseção é visualizar, através de mapas, o que já foi em parte mostrado graficamente na **subseção 3.3.2** e no **APÊNDICE 3.A**, p.155. Os arquivos com as divisões territoriais do Brasil foram obtidos no site do IBGE (http://www.ibge.gov.br/mapas_ibge/) e o *software* utilizado para colorir os mapas é o GeoDa (<https://www.geoda.uiuc.edu/>).

Nosso foco é o mercado de agências bancárias. Consideramos como não-submercado se a respectiva curva de Lorenz passa acima da curva-limite. Conforme visto na **Figura 3.2**, p.128 o Brasil todo não é um submercado. Este fato é visualizado graficamente no canto superior esquerdo da **Figura 3.10**, onde todo o território nacional aparece em cor única. Ao seu lado (direita superior), três estados na região norte aparecem em cor amarela, representando submercados (curva de Lorenz passa totalmente abaixo da curva-limite). À medida que aumentamos o nível de desagregação territorial, afloram proporcionalmente mais submercados. Em se tratando de mesorregiões, 42% delas representam um submercado. Se um determinado mercado apresenta 0 banco, não é possível afirmar com certeza se ele é ou não um submercado (ou se é parte de outro submercado). No caso das microrregiões, 2,5% delas se encaixam neste caso e por isso optamos por identificá-los em azul. Verifica-se também que grande área do território brasileiro é composta por um submercado (80% do total) quando consideramos as microrregiões (figura inferior direita).

A figura análoga para os municípios está no **APÊNDICE 3.D**, p.161 (**Figura A - 3.5**). Em somente 0,34% dos casos a curva de Lorenz passa acima da curva-limite e por isso a representação dos mesmos foi feita conjuntamente com os submercados. Quando passamos da configuração geográfica de microrregiões para municípios, a distribuição das agências bancárias varia bastante. Neste caso, 42% dos municípios não apresentam nenhum banco, que é visualizado pela considerável proporção de áreas azuis na figura. Cerca de 52% dos municípios são fragmentados totalmente, ou seja, cada banco participante possui apenas uma agência. Os 6% restantes também podem ser representados como um submercado. Para

melhor visualização, ampliamos as regiões NE, SE e S, respectivamente na **Figura A - 3.6**, **Figura A - 3.7** e **Figura A - 3.8**, p.162-163.

Figura 3.10 – Mapas de submercados – Brasil (esquerda superior), UFs (direita superior), mesorregiões (esquerda inferior) e microrregiões (direita inferior)

Legenda: ■ não-submercado ■ submercado ■ 0 banco

Fonte dos dados: Banco Central do Brasil.

As ilustrações apresentadas nesta subseção confirmam o resultado obtido ao longo deste capítulo para o mercado de agências bancárias. Se considerarmos as microrregiões, em 80% a curva de Lorenz passa abaixo da curva-limite. Em se tratando de municípios, exceto aqueles que não possuem nenhuma agência, quase todas as curvas de Lorenz passam abaixo da curva-limite. Do total de municípios, 52% são fragmentados totalmente. No entanto, à medida que se considera regiões geográficas mais amplas (mesorregiões, UFs e todo o território nacional) a assimetria no tamanho das firmas torna-se bastante elevada.

A consideração final que reforçamos sobre a utilização dos municípios como submercados é que o fato de que a maior parte deles possuir 0 ou 1 agências pode levar à distorção dos

resultados. Por exemplo, suponha que desejamos inferir sobre a relação entre participação e poder de mercado nesta indústria utilizando os dados municipais. Nos municípios com apenas uma agência bancária, o índice C1 será sempre igual a 100% e o Índice de Herfindahl, sempre igual a 1. Esses valores não indicam que o mercado é altamente concentrado, mas que o tamanho do mercado não suporta mais que um banco atuando. Por sua vez, municípios sem agências podem ser um submercado ou fazer parte de outros.

3.3.6 Tamanho das firmas

Nesta subseção buscamos mais evidências acerca da hipótese de agregação de submercados independentes de Sutton. A distribuição do tamanho das firmas apresenta bastante assimetria no Brasil (**Figura 3.2, subseção 3.3.2**, p.128) e nas UFs (**Figura 3.3, subseção 3.3.2**, p.129). Uma questão que emerge é se esse resultado não advém simplesmente da agregação de submercados onde as firmas pequenas são sempre pequenas. Existem firmas que atuam nacionalmente e possuem agências em quase todos os municípios e outras que atuam apenas localmente. Pretendemos mostrar que, dependendo do nível da agregação territorial, um banco que não ocupa as primeiras posições no *ranking* nacional pode ser grande em mercados locais.

Separamos os 118 bancos da nossa amostra em nacional e local, segundo o número de UFs com cobertura das agências. Consideramos como sendo de atuação nacional os seguintes bancos: Banco do Brasil, Caixa Econômica Federal, Bradesco, Itaú, ABN AMRO Real, HSBC e Unibanco. À exceção destes, todos os demais bancos foram classificados como locais. Apesar do número elevado, as agências do Banco Santander estão concentradas nas regiões Sul e Sudeste e, portanto este banco foi considerado de atuação local.

Para cada nível de agregação territorial (Brasil, UFs, mesorregiões, microrregiões e municípios) calculamos as respectivas participações de cada banco no mercado de depósitos, crédito e agências. Separamos os bancos em locais e nacionais e montamos o histograma, **condicionado** no número de mercados em que eles estão presentes (vide **Figura 3.11**, p.150 para o mercado de depósitos. Para os demais, vide **APÊNDICE 3.E**, p.164). Na **Figura 3.11**, o gráfico do canto superior direito (UFs) é interpretado da seguinte forma: dado que um banco local (histograma em vermelho) atua na UF, 95,5% possuem participação menor ou igual a

1% do mercado de depósito; 2,7% possuem entre 1-3%; 1,8% possuem entre 3-5%; e assim por diante. Os bancos nacionais aparecem em azul e possuem interpretação análoga.

Para a agregação do território nacional, verifica-se que a maioria (95,5%) dos bancos locais tem 1% ou menos de participação do mercado de depósitos. No entanto, à medida que desagregamos territorialmente, esse percentual se altera visivelmente. Nas UFs e nas mesorregiões, esse percentual cai para 86,3% e 70,4%, respectivamente. No caso dos municípios, em 21,6% das vezes, os bancos locais possuem 1% ou menos do mercado e em 14,5% das vezes, mais que 99%. A distribuição dos *market shares* dos bancos nacionais, por sua vez, não é afetada drasticamente pelo nível territorial. Na média, as firmas nacionais têm tamanho maior que as locais. No entanto, ambas apresentam distribuições similares para as microrregiões e municípios. Dessa forma, a desigualdade no tamanho dos bancos observada para o Brasil e para as UFs é decorrente do fato de que alguns têm atuação nacional e outros locais e não porque algumas firmas são sistematicamente grandes em todos os mercados. Resultado similar ocorre para os mercados de crédito e de agências, conforme observado na **Figura A - 3.9** e na **Figura A - 3.10** do **APÊNDICE 3.E**, p.164.

Figura 3.11 – Histograma da participação dos bancos no mercado de Depósito segundo a cobertura geográfica (Nacional ou Local)

Fonte dos dados: Banco Central do Brasil.

3.4 CONCLUSÃO

O objetivo deste capítulo é testar as predições do modelo de submercados independentes de Sutton no mercado de agências, depósito e crédito, utilizando os dados ESTBAN do Banco Central do Brasil de dez/2002.

A parte empírica do trabalho foi realizada da seguinte forma. Inicialmente, consideramos todo o território nacional e esboçamos as curvas de Lorenz. Para as três variáveis verifica-se que existe significativa desigualdade no tamanho dos bancos que operam no Brasil. Esse procedimento foi repetido para as UFs, mesorregiões e alguns municípios e constatou-se que à medida que aumentávamos o nível de desagregação territorial, as firmas passavam a apresentar tamanhos mais simétricos. Este primeiro resultado é consistente com a teoria de Sutton.

Com o intuito de identificar qual é a divisão geográfica que melhor representa um submercado, identificamos três delas para prosseguir com os testes: 1) as microrregiões; 2) os municípios; e 3) divisões geográficas distintas (mesorregiões, microrregiões, municípios e bairros) da região centro-oeste. Para os três casos, estimamos um *logit* ordenado no qual a variável dependente é o número de bancos que atua no submercado. Como variável independente utilizamos o número de agências e população tanto no submercado quanto nos submercados adjacentes e outros controles. Nossa hipótese é que se nossa amostra representa um submercado, então o número de bancos deve depender positivamente do número de agências dentro e negativamente (ou independe) das agências adjacentes. Os resultados obtidos sugerem que os três casos podem ser considerados submercados independentes. No caso dos municípios, parcela da amostra é composta por cidades com apenas um banco, indicando alta concentração e não gerando variabilidade nos dados. Por outro lado, o cômputo de diferentes divisões geográficas que representam um submercado representa um esforço muito grande por parte do pesquisador. Desta forma, adotar as microrregiões (sem considerar as mais populosas e as respectivas adjacências) como submercados independentes pode ser uma alternativa.

A identificação dos submercados independentes em municípios com mais de um submercado pode ser feita pela análise de cluster, onde diferentes clusters são identificados pela posição da

coordenada geográfica das agências. Utilizando esta técnica, para o município de Campo Grande identificamos seis submercados. Para estes submercados, as curvas de Lorenz para as agências bancárias passam abaixo da curva-limite: no entanto, o mesmo não ocorre para depósito e crédito. Interpretamos esse fato como sendo um indício de que o mercado de agências é geográfico e os de depósito e empréstimo, não abarca essa dimensão.

Para melhor visualização de como as regiões geográficas se comportam, esboçamos mapas (Brasil, UFs, mesorregiões, microrregiões e municípios) que ilustram os submercados. A mesma tendência mostrada graficamente pelas curvas de Lorenz é observada: o território nacional como um todo é pintado de uma única cor representando não-submercados e à medida que se desagrega territorialmente, essa cor vai perdendo área. Cerca de 80% das microrregiões podem representar um submercado. No caso dos municípios, a existência de diversos deles sem nenhum banco, bem como os submercados totalmente fragmentados, são facilmente visualizados no mapa.

Por fim, buscamos outras evidências nos dados que nos permitem testar as predições da teoria de Sutton. Neste caso, queremos nos certificar que a assimetria no tamanho das firmas observadas para o Brasil e para as UFs não advém do simples fato de agregar submercados em que as firmas pequenas são sempre pequenas. Separamos os bancos em dois grupos, sendo um composto por bancos de atuação nacional e outro, local. Verifica-se que quando se desagrega territorialmente, os bancos tentem a apresentar distribuição dos tamanhos mais similares.

De modo geral, este capítulo tenta entender melhor a estrutura do mercado bancário brasileiro, de modo que as discussões, os modelos e os testes a respeito da concentração sejam realizados dentro de submercados adequados.

3.5 REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHAM, Jean M.; GAYNOR, Martin S., VOGT, William B. *Entry and competition in local hospital markets*. **NBER Working Paper Series**. WP N°11.649. Sept., 2005.

BECK, Thorsten; DEMIRGÜÇ-KUNT, Asli; MAKSIMOVIC, Vojislav. *Bank competition and access to finance: international evidence*. **Journal of Money, Credit and Banking**. Vol. 36, N° 3. June, 2004.

BELAISCH, AGNÈS. *Do Brazilian banks compete?*. **IMF Working Paper WP/03/113**. 2003.

BRESNAHAN, Timothy F.; REISS, Peter C. *Entry in monopoly markets*. **The Review of Economic Studies**. Vol. 57, n° 4, pp. 531-553. Oct., 1990.

BRESNAHAN, Timothy F.; REISS, Peter C. *Entry and competition in concentrated markets*. **The Journal of Political Economy**. Vol. 99, n° 5, 977-1009. Oct., 1991.

BUZZACCHI, Luigi; VALLETTI, Tommaso M. *Firm size distribution: testing the “independent submarkets model” in the Italian motor insurance industry*. **International Journal of Industrial Organization**. Vol. 24, pp. 809-834. 2006.

CARDOSO, Marta Regina da Costa. **A concorrência na indústria bancária do Brasil**. Dissertação de Mestrado. FEA-USP. 2004.

CETORELLI, Nicola. *Entry and competition in highly concentrated markets*. **Economic Perspectives**. 4Q/2002

COELHO, Christiano A.; MELLO, João Manoel Pinho de; REZENDE, Leonardo. *Are public banks pro-competitive? Evidence from concentrated local markets in Brazil*. Working Paper. 2007. Disponível em: <http://www.econ.puc-rio.br/pdf/td551_revisto.pdf>. Acesso em 09/04/2008.

COHEN, Andrew; MAZZEO, Michael J. *Competition, product differentiation and quality provision: an empirical equilibrium analysis of bank branching decisions*. **Finance and Economics Discussion Series**. Federal Reserve Board, Washington, D.C. - 46. Jun., 2004.

DE JUAN, Rebecca. *The independent submarkets model: an application to the Spanish retail banking market*. **International Journal of Industrial Organization**. Vol. 21, pp. 1461-1487. 2003.

DICK, Astrid A. *Market size, service quality, and competition in banking*. **Journal of Money, Credit and Banking**. Vol. 39, n° 1. Feb., 2007.

GIANNETTI, Caterina. *Intensity of competition and market structure in the Italian banking industry*. Discussion Paper N°2008-43. **Tilburg University**. April, 2008.

LONG, J. Scott; FREESE, Jeremy. *Regression models for categorical dependent variables using Stata*. STATA Press. 2nd Edition, 2006.

NAKANE, Márcio I. *A test of competition in Brazilian banking*. **Banco Central do Brasil**. Working Paper Series n. 12. Março, 2001.

NAKANE, Márcio I.; ALENCAR, Leonardo S. e KANCZUK, Fábio. Demanda por serviços bancários. In: FIUZA, Eduardo P.S.; MOTTA, Ronaldo Seroa da (Coord. Téc.). **Métodos quantitativos em defesa da concorrência e regulação econômica**. Tomo 1. Rio de Janeiro, IPEA, 2006.

ROCHA, Fernando Alberto Sampaio. Evolução da concentração bancária no Brasil (1994-2000). **Notas Técnicas do Banco Central do Brasil**. N° 11. Nov., 2001.

RABE-HESKETH, Sophia; EVERITT, Brian. *A handbook of statistical analyses using Stata*. 3rd Edition. Chapman & Hall/CRC. 2004.

SUTTON, John. *Technology and market structure: theory and history*. The MIT Press. 1998.

SUTTON, John. *Market Structure: Theory and Evidence*. In: ARMSTRONG, Mark; PORTER, Robert H. (Editors). *Handbook of industrial organization*. Vol. 3. 2007.

WALSH, Patrick Paul; WHELAN, Ciara. *Portfolio effects and firm size distribution: carbonated soft drinks*. **The Economic and Social Review**. Vol. 33, n° 1, pp. 43-54. Spring, 2002.

WOOLDRIDGE, Jeffrey M. *Econometric analysis of cross section and panel data*. The MIT Press. 2002.

ZHOU, Xiaolan. *Estimation of the impact of mergers in the banking industry*. Job Market Paper. Apr., 2008. Disponível em: <http://pantheon.yale.edu/~xz55/Zhou_bank>. Acesso em 09/04/2008.

APÊNDICE 3.A

Figura A - 3.1 – Curva de Lorenz do volume de depósito das Mesorregiões – Região Norte e demais UFs
A curva pontilhada fina refere-se às mesorregiões. A curva preta pontilhada grossa é a curva-limite. Em verde contínuo, a curva da UF (não-reportado para a região norte).

Fonte dos dados: Banco Central do Brasil.

Figura A - 3.2 – Curva de Lorenz do volume de crédito das Mesorregiões – Região Norte e demais UF's
 A curva pontilhada fina refere-se às mesorregiões. A curva preta pontilhada grossa é a curva-limite. Em azul contínuo, a curva da UF (não-reportado para a região norte)

Fonte dos dados: Banco Central do Brasil.

Figura A - 3.3 – Curva de Lorenz de agências das Mesorregiões – Região Norte e demais UF's

A curva pontilhada fina refere-se às mesorregiões. A curva preta pontilhada grossa é a curva-limite. Em vermelho contínuo, a curva da UF (não-reportado para a região norte).

Fonte dos dados: Banco Central do Brasil.

APÊNDICE 3.B

Tabela A - 3.1 – Resultados da estimativa do *generalized ordered logit model* – municípios (2)

Variáveis	Categorias					
	1	2	3	4	5	6
<i>nbranch</i>	53,264 * 3,526	24,557 * 1,388	26,212 * 2,178	7,199 * 1,076	4,691 * 0,874	1,501 * 0,262
<i>adbranch</i>	0,101 & 0,042	0,040 0,056	-0,004 0,018	-0,058 # 0,032	0,003 0,020	-0,026 0,026
<i>lpopt</i>	-5,069 * 1,719	-1,398 1,011	-2,920 & 1,378	-1,413 # 0,808	-0,341 0,897	-0,114 0,768
<i>ladpopt</i>	1,648 1,418	-0,435 0,363	-0,565 0,616	0,303 0,190	-0,181 0,217	0,026 0,203
<i>dens</i>	0,054 * 0,011	0,000 0,007	0,016 0,013	0,006 0,004	-0,007 0,005	0,001 0,001
<i>hdep</i>	-228,390 * 13,454	-16,510 * 3,258	-8,619 # 4,564	-21,311 * 6,643	-25,593 * 6,928	-20,129 * 6,461
<i>area</i>	-0,106 # 0,063	-0,056 0,038	-0,096 * 0,023	-0,038 * 0,010	-0,016 0,010	0,037 0,117
<i>dummy</i>	28,989 * 1,839	0,955 1,276	0,643 0,821	1,239 & 0,543	1,314 * 0,700	5,156 * 1,590
<i>constante</i>	114,244 * 10,087	-56,402 * 5,437	-86,134 * 4,212	-24,000 * 5,327	-18,961 * 4,252	-12,886 * 3,275

Nº observações: 2.757. Desvio-padrão robusto à heterocedasticidade.

Redefiniu-se $nbanks^* = 6$ se $nbanks^* = 7$ e $nbanks^* = 7$ se $nbanks^* \geq 8$.

Dummy igual a 1 para as regiões SE e S e zero caso contrário.

*, & e # indicam coeficiente da variável significativo a 1%, 5% e 10%, respectivamente.

2.233 probabilidades negativas preditas. Estimação usando o comando **gologit2**.

Tabela A - 3.2 – Resultados da estimativa do *generalized ordered logit model* – municípios (3)

Variáveis	Categorias				
	2	3	4	5	6
<i>nbranch</i>	38,812 * 5,401	26,421 * 1,584	7,106 * 1,007	4,697 * 0,799	1,503 * 0,256
<i>adbranch</i>	0,003 0,045	-0,030 0,027	-0,058 # 0,034	-0,008 0,025	-0,026 0,026
<i>lpopt</i>	-0,976 1,377	-1,958 1,833	-1,487 0,964	-0,003 0,826	-0,118 0,734
<i>ladpopt</i>	-0,557 0,565	-0,263 0,222	0,209 0,187	-0,038 0,278	0,024 0,210
<i>dens</i>	0,018 0,016	0,014 0,013	0,010 * 0,003	-0,006 0,006	0,001 0,001
<i>hdep</i>	-24,674 * 7,164	-17,578 # 9,243	-25,190 * 6,777	-37,478 * 9,291	-20,086 * 6,667
<i>area</i>	-0,105 0,078	-0,150 0,161	-0,042 * 0,009	-0,019 # 0,010	0,038 0,110
<i>loancv</i>	5,122 * 1,256	5,576 # 3,233	3,402 & 1,400	3,055 * 0,889	-0,052 0,994
<i>dummy</i>	1,494 1,359	1,816 1,785	1,140 # 0,598	1,898 & 0,865	5,134 * 1,717
<i>constante</i>	-100,605 * 15,934	-92,992 * 8,258	-24,978 * 6,498	-20,998 * 3,823	-12,821 * 3,514

Nº observações: 1.487 (exclui $nbanks^* = 0$). Desvio-padrão-robusto à heterocedasticidade.

Redefiniu-se $nbanks^* = 6$ se $nbanks^* = 7$ e $nbanks^* = 7$ se $nbanks^* \geq 8$.

Dummy igual a 1 para as regiões SE e S e zero caso contrário.

*, & e # indicam coeficiente da variável significativo a 1%, 5% e 10%, respectivamente.

1.007 probabilidades negativas preditas. Estimação usando o comando **gologit2**.

APÊNDICE 3.C

Figura A - 3.4 – Submercados de Campo Grande - MS

Fonte: Google Maps (<http://maps.google.com/>).

APÊNDICE 3.D

Figura A - 3.5 – Mapas de submercado – Municípios

Legenda: não-submercado e submercado totalmente fragmentado 0 banco

Fonte: Banco Central do Brasil.

Figura A - 3.6 – Mapas de submercado – Municípios da região NE

Legenda: não-submercado e submercado totalmente fragmentado 0 banco

Fonte: Banco Central do Brasil.

Figura A - 3.7 – Mapas de submercado – Municípios da região SE

Legenda: não-submercado e submercado totalmente fragmentado 0 banco

Fonte: Banco Central do Brasil.

Figura A - 3.8 – Mapas de submercado – Municípios da região S

Legenda: ■ não-submercado e submercado ■ totalmente fragmentado ■ 0 banco

Fonte: Banco Central do Brasil.

APÊNDICE 3.E

Figura A - 3.9 – Histograma da participação dos bancos no mercado de Crédito segundo a cobertura geográfica (Nacional ou Local)

Fonte dos dados: Banco Central do Brasil.

Figura A - 3.10 – Histograma da participação dos bancos no mercado de Agências segundo a cobertura geográfica (Nacional ou Local)

Fonte dos dados: Banco Central do Brasil.