

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

Vizinhos, vizinhos, negócios à parte.

O efeito da proximidade entre agentes e clientes em contratos de microcrédito.

Alexandre Spitz

No. de Matrícula: 0511504

Orientador: Juliano Assunção

Co-orientadora: Isabelle Agier

Dezembro, 2009

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

Vizinhos, vizinhos, negócios à parte.

O efeito da proximidade entre agentes e clientes em contratos de microcrédito.

Alexandre Spitz

No. de Matrícula: 0511504

Orientador: Juliano Assunção

Co-orientador: Isabelle Agier

Dezembro, 2009

Declaro que o presente trabalho é de minha autoria e que não recorri para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor.

Aos meu pais, por me ensinarem a pensar o mundo e escolher caminhos,

e à Liz, por estar presente em meus pensamentos e escolhas.

Agradecimentos especiais a Isabelle Agier, pela orientação e contribuições para este estudo, e Juliano Assunção por ajudar a solucionar questões durante meu percurso.

Agradeço também a toda a equipe do VivaCred, em especial Teófilo Cavalcanti, que autorizou a utilização da base de dados.

Sumário

1) Introdução:	5
2) Microcrédito e informação assimétrica.	8
3) VivaCred	14
4) Base de dados:	18
A) Base de Endereços.....	18
B) Base de Contratos.....	29
C) Base de Contratos e Distâncias	31
5) Resultados Empíricos:	36
6) Conclusão:	39
7) Bibliografia:	40

1) Introdução:

Esta monografia analisa - através de evidência empírica - a relevância do conhecimento local para o desempenho de contratos de microcrédito. Regredindo-se a distância entre a residência de um agente de crédito e a de seus clientes em uma variável *dummy* de atraso da amortização, verificou-se que quanto mais distantes entre si são as moradias, maior a probabilidade de atraso. Mais precisamente, as regressões sugerem que quando se dobra a distância, a probabilidade de se ter um dos pagamentos fora do prazo aumenta em 0.3 pontos percentuais, e é estatisticamente significativa. Considerando-se que houve atrasos em pelo menos uma das parcelas da amortização em 8.6% dos empréstimos, este coeficiente é também economicamente relevante.

Para este estudo foram utilizados dados provenientes do VivaCred, instituição de microcrédito sem fins lucrativos, desde seu início em 1997 até 2007. A base abrange todas as seis agências: Rocinha, Rio das Pedras, Maré, Santa Cruz, Glória e Macaé. Estão cadastradas nesta base 35.194 pessoas que demandaram empréstimos e 31.691 empréstimos concedidos. Dentre os que demandaram empréstimos, somente 11.260 tiveram seus créditos aprovados, sendo que parte deles mais de uma vez. Durante estes onze anos de operação, passaram pelos quadros da instituição 40 analistas, sendo que em torno de 20 estão em atividade, em média 3 por agência.

O processo para se chegar às regressões teve início em um intenso trabalho de limpeza, padronização e formatação dos endereços dos clientes e analistas. Isto foi necessário para poder encontrar as coordenadas relativas a cada endereço e calcular as distâncias entre eles. Para esta transformação foi utilizada a ferramenta de geocodificação² do Google Maps. Como grande parte dos clientes do VivaCred reside em favelas, seus endereços não foram reconhecidos por este método. Foi então necessário desenvolver métodos alternativos de localização dos clientes por seus subbairros - regiões definidas como um subconjunto de um bairro - especialmente para favelas.

Agentes de crédito participam ativamente tanto na concessão do crédito quanto no monitoramento de seu desempenho. O resultado acima indica que a intermediação desse processo por um agente³ conhecedor de seu ambiente de trabalho é uma solução

² Transformação de endereços em coordenadas geográficas.

³ Também chamado de “analista de crédito” ao longo desta monografia.

eficiente para se mitigar o efeito negativo de informações assimétricas. Confirma-se assim a hipótese inicial desta pesquisa de que quanto mais próximo o analista estiver de seus clientes, melhor este poderá gerenciar tanto a oferta de novos contratos quanto o repagamento dos que estão ativos.

O momento em que está sendo discutida a possibilidade de empréstimo entre o agente e um cliente em potencial é essencial para o primeiro ter clareza sobre as condições de pagamento do segundo, evitando-se desta forma uma seleção adversa. Nessa conversa são feitas pelo agente diversas perguntas a fim de obter informações sobre quem é o tomador do empréstimo e qual a situação financeira de sua família e de seu negócio. Conhecer a realidade da pessoa que se está questionando é uma poderosa ferramenta de trabalho para o analista, possibilitando a interpretação correta das informações que lhe estão sendo passadas. Além disso, vivenciar o dia-a-dia da localidade onde trabalha permite ao agente desenvolver intuição sobre a verdadeira capacidade de repagamento dos clientes.

Uma vez concedido o empréstimo, o papel do agente passa a ser o de monitorar o desempenho desses contratos, ou seja, verificar se as parcelas da amortização estão sendo pagas em dia. Isto é especialmente importante para os analistas em virtude de sua influência na parte variável de seus salários: caso alguma parcela de um empréstimo concedido atrase mais de 30 dias, seu salário será menor. Para que isto não aconteça, o agente realiza diariamente uma rotina de visitas a clientes que perderam o prazo de pagamento por um ou dois dias, e tenta garantir que a parcela será paga. Em especial para o microcrédito, pelo caráter informal das relações, residir próximo de seus clientes permite ao analista minimizar o risco moral, observando cotidianamente as atividades destes e o andamento de seus negócios.

Esta monografia se mostra especialmente relevante no momento em que a economia mundial, após a grande crise dos últimos anos, começa a se reestruturar. Economistas divergem sobre os motivos exatos que levaram o mundo à “crise do *subprime*”, porém, como demonstra Mishkin (2009), há um consenso de que esta foi essencialmente causada devida à informação assimétrica. O ciclo de crédito de hipotecas nos Estados Unidos se tornou tão complexo que os investidores no final da cadeia não tinham mais conhecimento sobre a real condição do ativo que estavam comprando. Os agentes de crédito, por sua vez, conheciam as pessoas para quem estavam emprestando, e suas possíveis incapacidades de repagamento, mas ignoravam

esta questão ao verificar a incessante demanda por estes ativos, e conseqüentemente a renda proveniente dessas vendas.

A conexão entre a oferta e demanda por estas hipotecas de alto risco se fez por instituições que agrupam blocos de empréstimos, processo conhecido como “securitização”, e vendem estes aos investidores. A estes blocos de empréstimos foram conferidas notas de risco de crédito (*rating*) por agências de classificação, de acordo com padrões estatísticos pré-definidos. Dessa maneira, quem comprava estes ativos inferia o risco por esta nota, fazendo com que gradualmente fosse eliminada da cadeia a informação proveniente do agente de crédito sobre como era percebido o cliente em relação a sua capacidade de amortização de dívidas. Uma vez identificado o problema, entendeu-se que o conhecimento local dos agente é muito relevante para e não pode ser substituído por métodos matemáticos.

Antes desse, outros estudos já analisaram métodos institucionais de se mitigar o efeito de informações assimétricas em contratos de microcrédito. Na maioria dos casos, como em Ghatak e Guinane (1999), o método analisado é o empréstimo a grupos, em que se diversifica o risco.

Esta monografia amplia o entendimento sobre a relevância do agente de crédito para o microcrédito, em especial para o VivaCred, continuando o estudo realizado por Agier e Assunção (2009)⁴ denominado “*The role of credit officers in the performance of microcredit loans: evidence from VivaCred in Brazil*”.

⁴ Este trabalho não seria possível sem as contribuições de Isabelle Agier e Juliano Assunção. Desde a construção da base de dados até discussões sobre a escolha de amostras, estes tiveram papel essencial nesta monografia.

2) Microcrédito e informação assimétrica.

Tido por muitos como um dos mais relevantes instrumentos de desenvolvimento social, o microcrédito estabelece uma ponte entre empreendedores de baixa renda - antes excluídos de qualquer possibilidade de tomar empréstimos - e o mercado financeiro.

Tradicionalmente destinado a indivíduos com baixo poder aquisitivo, o microcrédito era visto como uma forma de assistencialismo. Entretanto, nos últimos anos vem ganhando gradativamente espaço no debate de finanças, e começa-se a enxergá-lo como um produto que pode ser, ao mesmo tempo, altamente rentável e socialmente produtivo.

Um recente relatório do Deutsche Bank intitulado “*Microfinance: an emerging investment opportunity*” chama a atenção exatamente para este fenômeno – denominado de “resultado duplo” (*double bottom line*) - e destaca a crescente demanda por este tipo de investimento, tanto pela necessidade de capital das populações de baixa renda, quanto por investidores em busca de altas taxas de retorno. No Brasil, desde 2005, ano da criação do Programa Nacional de Microcrédito Produtivo Orientado (PNMPO) – este modelo de financiamento já dobrou de tamanho, como visto na Tabela 1.

Estes pequenos empréstimos tornaram-se um grande fenômeno, recebendo cada vez mais atenção e recursos, obtendo rápido crescimento tanto em termos de alcance quanto em termos do montante de dinheiro emprestado. Conseqüentemente, mais pessoas podem contar com os benefícios da alocação intertemporal de capital, ou seja, podem transformar renda futura em consumo presente, e passam a poder transformar suas idéias e dedicação em negócios.

Tabela 1: Crescimento do Microcrédito no Brasil

Ano	Operações de Microcrédito (em unidade)	Índice ¹	Valor Concedido ² (em R\$ 1,00)	Índice ³
2005	632.106	100	602.340.000,00	100
2006	828.847	131,12	831.815.600,78	138,10
2007	963.459	152,42	1.100.375.829,94	182,68
2008	1.274.296	201,60	1.807.071.717,91	300,01
2009*	347.761	-	479.752.458,84	-
Total	4.046.469		4.821.355.607,47	

Fonte: PNMPO, 2009.

Notas:

* Dados referentes ao 1º. Trimestre de 2009

¹ Índice da Quantidade de Operações de Microcrédito Produtivo Orientado. Ano Base 2005 = 100,00.

² Valores Nominais.

³ Índice do Volume de Crédito Concedido para o Microcrédito Produtivo Orientado. Ano Base 2005 = 100,00.

O formato que conhecemos hoje ganhou visibilidade em Bangladesh, na década de 1970, quando Muhammad Yunus⁵ concedeu um pequeno empréstimo a agricultores que precisavam de capital para iniciar suas culturas, mas que, sem condição de oferecer garantia alguma, não podiam recorrer aos bancos. Os bancos, por sua vez, não conseguiam aferir corretamente o risco de *default*⁶ que estas pessoas representavam.

Mercados de crédito podem não funcionar corretamente devido, principalmente, à assimetria de informação, ou seja, quando as duas partes de uma negociação ou contrato têm entendimentos diferentes sobre ele. Este fenômeno foi inicialmente estudado por Akerlof (1970)⁷ em artigo denominado “*The Market for Lemons: Quality Uncertainty and the Market Mechanism*”, no qual focaliza o caso do mercado de compra e venda de carros usados. Akerlof inicia seu artigo da seguinte maneira:

“Este artigo relaciona qualidade e incerteza. A existência de bens de muitas qualidades diferentes impõe problemas interessantes e importantes para a teoria de mercado.” (Akerlof 1970: 488)

Trazendo esta relação para o mercado de microcrédito, e olhando sob a perspectiva da firma, entendemos “*bens*” como os tomadores de empréstimos e “*qualidade*” como a propensão destes a repagar o empréstimo de acordo com o contrato. Esta incerteza sobre a qualidade pode aparecer em dois momentos distintos: tanto quando o empréstimo está sendo negociado, quanto ao longo de seu repagamento. A figura 1 resume a situação descrita a seguir.

No momento inicial de um contrato, é muito importante que o emprestador tenha clareza sobre quem é, de fato, seu cliente e qual sua capacidade de repagamento, para que seja feita a seleção dos melhores clientes. Segundo Akerlof, existe uma tendência a haver uma “seleção adversa”, o que quer dizer que há maiores chances de serem selecionados os clientes com menor “qualidade”. Dado que o preço do empréstimo – a

⁵ Fundador e Presidente do Banco Grameen e com este, Prêmio Nobel da Paz de 2006

⁶ Default é o termo usado para caracterizar o não cumprimento de uma cláusula de um contrato, tal como o não pagamento de uma das parcelas.

⁷ George A. Akerlof recebeu, junto com Michael Spence e Joseph Stiglitz, o Prêmio Nobel de Economia de 2001 por suas análises de mercados com assimetria de informação.

taxa de juros – é a mesma⁸, possivelmente a maior parte da demanda virá dos indivíduos com maior propensão ao risco. Isto ocorre pois risco e retorno estão positivamente correlacionados, fazendo com que os clientes que esperam ter os maiores retornos também sejam os principais candidatos a *default* em seus empréstimos.

Por este motivo, quando se encontram o demandante e o ofertante para discutir a possibilidade e os termos de um empréstimo, enquanto o primeiro está tentando demonstrar que é um bom pagador, o segundo tenta verificar se este fala a verdade. Estes dois lados da negociação ficaram respectivamente conhecidos como sinalização (signaling) - estudado inicialmente por Spence (1973)⁹ - e triagem (*screening*) - estudado inicialmente por Stiglitz (1975)¹⁰.

No segundo momento de um contrato, quando o cliente já obteve o empréstimo e está realizando sua amortização, novas incertezas surgem e são provenientes da falta de conhecimento por parte do prestador sobre o que está sendo feito com o empréstimo, o que pode ser chamado “risco moral” (*moral hazard*). Este foi inicialmente estudado em Arrow (1963)¹¹, e pode ser dividido em dois tipos: involuntário ou estratégico.

O tipo involuntário de risco moral é a dificuldade de se monitorar a real finalidade do empréstimo. Um dinheiro que foi tomado para a compra de insumos pode estar sendo usado em uma atividade de maior risco do que inicialmente se pensava, ou, por exemplo, para pagar uma festa, prejudicando a habilidade de pagamento do tomador. Isto pode levar à não capacidade de se repagar o empréstimo, o que ficaria caracterizado como um *default* involuntário.

Um exemplo clássico deste problema vem do crédito a grupos de agricultores. Se um grande banco, que não conhece bem os riscos associados às atividades destes, resolve conceder-lhes um empréstimo, pode ter problemas com a verificação do que de fato aconteceu no caso do não repagamento. Questões como falta de chuvas, seca ou pragas podem ser facilmente verificadas, mas acentuam a correlação entre a produção de fazendas próximas, consequentemente levando à correlação também da receita destes agricultores. O banco pode tentar minimizar esse problema emprestando a grupos em

⁸ O que veremos ser verdade mais adiante ao analisar-se o VivaCred, instituição que serviu de base para os estudos empíricos.

⁹ “Job Market Signaling”, 1973

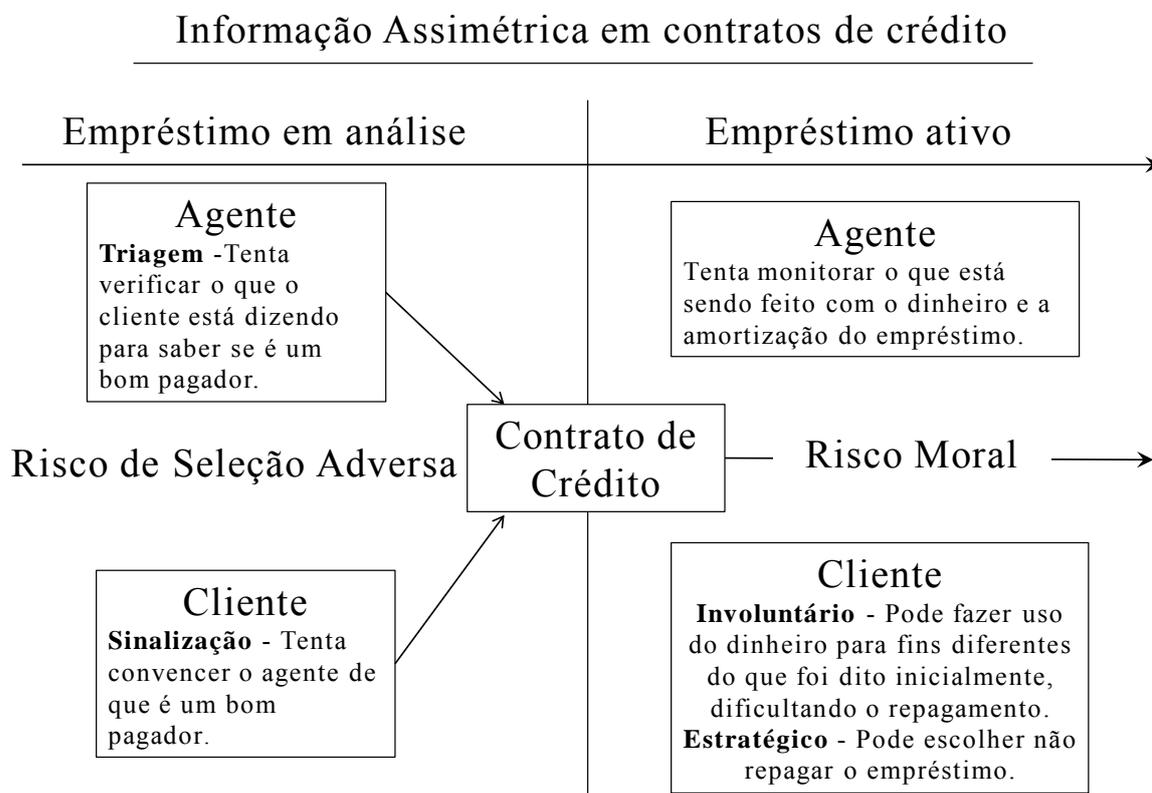
¹⁰ “The Theory of “Screening,” Education, and the Distribution of Income”, 1975

¹¹ “Uncertainty and the Welfare Economics of Medical Care”, 1963

uma área maior, porem com isso aumentará sua dificuldade de aferir outro fator tão importante quanto o efeito da natureza, e bem mais difícil de ser mensurado: o esforço que cada indivíduo empenhou em sua colheita, com o objetivo de honrar seu compromisso com a instituição que lhe concedeu o empréstimo.

O tipo estratégico de risco moral se apresenta quando um cliente emprega o dinheiro exatamente no que disse que ia fazer, mas escolhe não repagar o empréstimo, o que seria um *default* voluntário ou estratégico. Isto pode acontecer especialmente em ambientes de negócios onde as leis não são aplicadas com a severidade necessária, tornando muito difícil a punição do mau pagador, como acontece em países em desenvolvimento.

Figura 1: Informação assimétrica em contratos de crédito



Constata-se assim que existem diversos problemas de informação assimétrica na criação e manutenção de contratos de crédito. Para estes, foram pensadas diversas soluções, tanto por acadêmicos quanto por instituições bancárias. Dentre elas, a que se tornou padrão para os mercados financeiros foi a de requerimento de colateral: para a instituição que está emprestando, os problemas acima citados perdem relevância quando se tem a segurança de ativos penhorados em seu favor.

Mesmo se sua seleção foi adversa, ou se o tomador fez mau uso do dinheiro - caso o cliente não repague corretamente seu empréstimo a instituição tem o direito de recorrer à justiça para que lhe sejam transferidos estes ativos, fazendo com que o risco de perdas significativas tenda a zero.

Entretanto, emprestadores só aceitam como colateral ativos líquidos e que tenham valor pelo menos semelhante ao montante do empréstimo, fazendo com que esta solução não atenda às necessidades das instituições de microfinanças. Por terem como clientes principalmente indivíduos pertencentes a camadas de baixa renda, torna-se impraticável para tais instituições exigir de seus clientes algum tipo de colateral que atenda a essas especificações. Muhammad Yunus, em uma cartilha sobre microcrédito para empreendedores de 1986, escreve:

“Argumentar que não se pode emprestar aos pobres porque estes não têm colateral é o mesmo que argumentar que o homem não pode voar porque não tem asas.” (Yunus, M. 1986)

A criação de alternativas no sentido de minimizar este problema foi exatamente o que tornou o microcrédito tão relevante. Dentre elas, duas alternativas se tornaram as mais utilizadas: empréstimos a grupos (*group lending*) e o agente de crédito (*credit officer*).

A história dos empréstimos a grupos se confunde com a história do microcrédito em si. Desde o início do Grameen Bank, em Bangladesh, esta foi a maneira pensada para se lidar com empréstimos em que o tomador não oferecia garantias. Unindo-se os tomadores do empréstimo em grupos, diversifica-se o risco, diminuindo as chances de que o grupo não consiga repagar o crédito. Caso um dos participantes do grupo não consiga cumprir com suas obrigações em um determinado mês, o grupo as cumpre por ele e posteriormente este cobrirá suas dívidas.

Paralelamente ao desenvolvimento do crédito a grupos, pensou-se em uma maneira de minimizar os problemas advindos da falta de informação sobre o cliente, o agente de crédito. Este seria encarregado de fazer a intermediação entre a instituição e a sociedade, tanto na hora da seleção de clientes quanto no acompanhamento dos empréstimos destes.

Assim, em diferentes países e instituições o microcrédito apresenta diferentes nuances¹², de forma a melhor se adaptar às diversass culturas e necessidades. No Brasil, optou-se pelo microcrédito produtivo orientado como o modelo a ser estimulado pelo governo, e foi definido da seguinte maneira:

“(...) considera-se microcrédito produtivo orientado o crédito concedido para o atendimento das necessidades financeiras de pessoas físicas e jurídicas empreendedoras de atividades produtivas de pequeno porte, utilizando metodologia baseada no relacionamento direto com os empreendedores no local onde é executada a atividade econômica (...)” (Art. 1º, § 3 da Lei Nº 11.110, de 25 de Abril de 2005).

Neste mesmo parágrafo da lei lê-se ainda:

“O atendimento ao tomador final dos recursos deve ser feito por pessoas treinadas para efetuar o levantamento socioeconômico e prestar orientação educativa sobre o planejamento do negócio, para definição das necessidades de crédito e de gestão voltadas para o desenvolvimento do empreendimento;” (Art. 1º, § 3 da Lei Nº 11.110, de 25 de Abril de 2005).

Esta monografia se concentra no estudo do papel dessas pessoas, conhecidas como analistas de crédito, essenciais ao funcionamento do sistema e que intermedeiam os empréstimos entre as instituições prestadoras e seus clientes. Em especial, busca-se entender o quão relevante é, para o desempenho do empréstimo, o entendimento, pelo analista, da realidade local de seu ambiente de trabalho. Ou seja, em que medida conhecer profundamente quem são esses clientes e seu dia-a-dia afeta a qualidade desses micro-empréstimos.

¹² Podem ser variações no modelo escolhido: com ou sem direcionamento específico, produção ou consumo, individual ou coletivo, urbano ou rural, o tamanho dos empréstimos e se são exclusivamente para mulheres.

3) VivaCred

O VivaCred é uma instituição sem fins lucrativos que realiza microcréditos à população de baixa renda do Estado do Rio de Janeiro, principalmente em favelas. Sua missão é:

“Proporcionar condições de acesso ao crédito aos micro e pequenos empreendedores, especialmente nas comunidades carentes, do setor formal ou informal da economia, buscando desse modo promover o desenvolvimento econômico e a integração social no Rio de Janeiro.” (www.vivacred.com.br)

A primeira agência foi inaugurada na favela da Rocinha no final de 1996. Desde então, cinco outras agências foram criadas: Rio das Pedras em 1998, Copacabana (agora na Glória) em 1999, Favela da Maré em 2000, Santa Cruz em 2002 e na cidade de Macaé em 2004.

Desde o início até o final de 2008, seu maior financiador foi o Banco Nacional de Desenvolvimento Econômico e Social (BNDES), permitindo que, com subsídio governamental, fossem cobradas baixas taxas de juros. A partir do início de 2009 as operações do VivaCred foram fundidas com as do CrediAmigo do Banco do Nordeste. Assim, este passou a ser o financiador e controlador operacional da instituição. Nesta monografia estão sendo analisados dados de empréstimos de 1997 a 2007, fazendo com que esta mudança de controle não tenha efeito.

Dentro da estrutura do VivaCred, os agentes de crédito tem um importante papel na redução de problemas advindos de informação assimétrica. Este tem duas tarefas principais durante o processo, sendo a primeira enquanto o contrato está sendo negociado e a segunda quando o contrato está ativo:

- (1) Analisar a capacidade de pagamento de um potencial cliente e apresentá-las ao comitê de crédito para aprovação;
- (2) Verificar a situação de cada contrato regularmente e impedir que as amortizações atrasem.

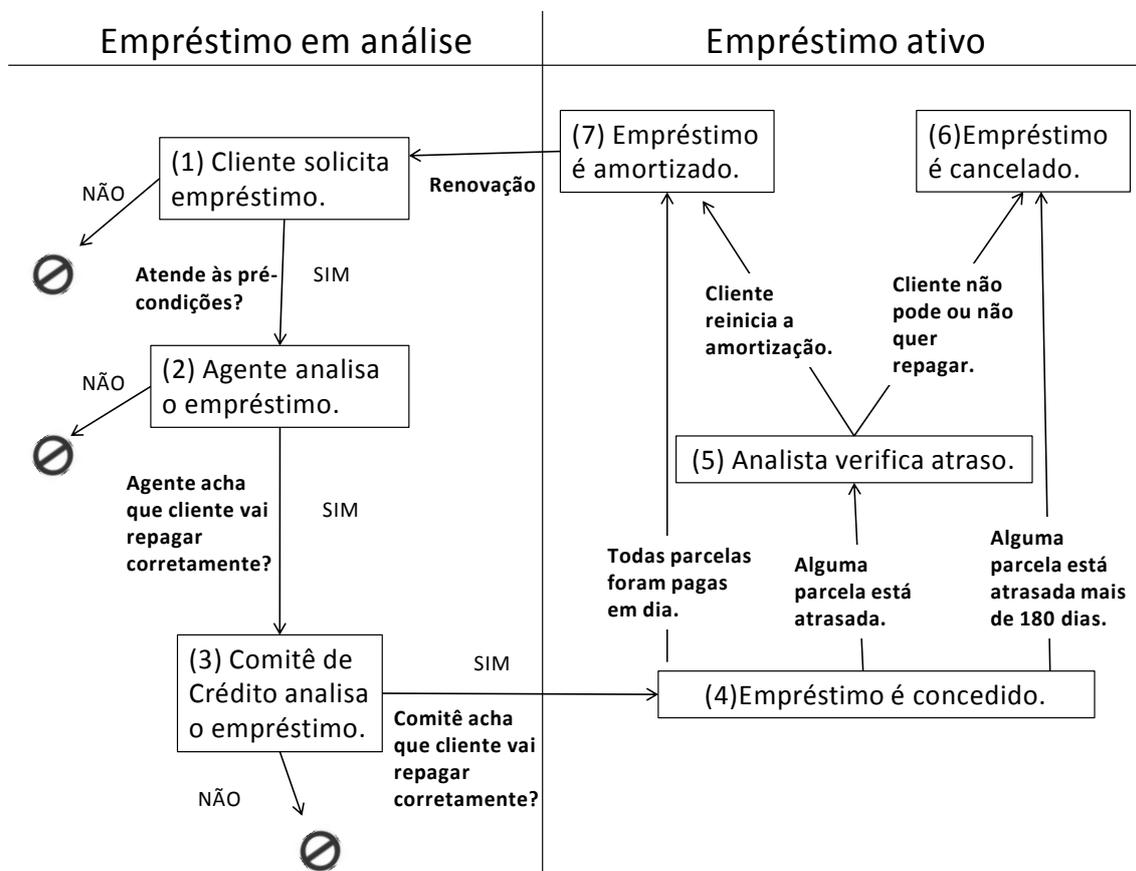
É importante entender como é o ciclo de um contrato para melhor entender as partes do processo que são influenciadas pelo agente de crédito, como pode ser visto na Figura 2.

De acordo com Agier e Assunção (2009) os passos podem ser definidos da seguinte forma:

“O processo tem início (passo 1) quando um cliente faz o pedido de empréstimo. Para ser elegível, o negócio para onde está indo o dinheiro tem que ter sido criado com mais de 6 meses de antecedência e o nome do cliente não pode constar na lista de devedores do Serviço de Proteção ao Crédito (SPC). Para fazer uma solicitação o cliente tem que apresentar documentos de identificação dele e de seu fiador, assim como documentos do negócio.” (Agier e Assunção 2009)

Por mais que isto pareça um detalhe no processo, por se tratar de populações de baixa renda e localidades não reconhecidas oficialmente, documentação pode ser um problema. Em localidades como a Favela da Rocinha, o endereço do cliente é reconhecido através de um documento da Associação de Moradores da Rocinha, afirmando que o cliente mora na rua que diz morar. Além disso, o propósito do empréstimo deve ser explicado.

Figura 2: Ciclo de crédito no VivaCred



Fonte: Agier e Assunção (2009)

Uma questão importante a ser entendida no âmbito desse estudo é a distribuição dos agentes. Cada agência conta em média com 3 agentes que se dividem por área. Cada um deles é responsável por realizar todas as partes do processo junto a cada um dos clientes da sua área. Trabalhar todos os dias na mesma área permite uma maior interação do agente com seus clientes, e conseqüentemente, mais conhecimento de suas capacidades de repagamento de novos empréstimos e o que está sendo feito com empréstimos já concedidos.

“O agente de crédito visita o cliente e o fiador (passo 2) para coletar as informações sobre colateral, solvência e risco do negócio. O agente encontra o possível cliente em seu estabelecimento. É pedido ao cliente que preencha um questionário sobre sua situação pessoal, seu orçamento familiar e a situação financeira de seu negócio. As mesmas perguntas são feitas para o fiador, entretanto com menos detalhes para as informações financeiras” (Agier e Assunção 2009)

O primeiro momento em que se encontram o analista e cliente é de essencial importância para o bom desempenho do empréstimo. A maioria dos clientes não está acostumada a responder o tipo de pergunta que lhes são feitas, exigindo do analista habilidade em conduzir a conversa. É neste momento que faz diferença o fator sobre o qual esta monografia se dedica a estudar, o conhecimento local. Conhecendo quem são, ou pelo menos como vivem e trabalham estes possíveis clientes, o analista sabe fazer as perguntas certas para obter as respostas que procura, assim como interpretar o que lhe é dito. Um exemplo de técnica muito utilizada pelos analistas é fazer a mesma pergunta em dois momentos distintos de uma conversa e perceber se a resposta dada foi a mesma, mostrando consistência.

Após esta conversa, a aplicação, é levada ao comitê de crédito (passo 3) junto com as observações pessoais do analista. O comitê tem a palavra final sobre a concessão de um empréstimo, assim como seus termos. Caso seja aprovado, dinheiro é liberado (passo 4) e cliente inicia a amortização no mês seguinte. Empréstimos podem ser pagos mensal ou quinzenalmente.

Continuando o ciclo de crédito, conforme Agier e Assunção (2009), depois que o dinheiro é liberado o empréstimo pode tomar 3 caminhos:

“Em caso de amortização completa, sem atrasos significantes (passo 7), o cliente tem acesso a outro empréstimo, possivelmente com valor maior. Se não houve atrasos durante todo o período do contrato, a taxa de inscrição é reduzida.

Em caso de atraso (passo 5), os agentes cumprem outra função importante no processo. Isto se inicia com o agente visitando cliente e tentando convencer este a pagar a parcela devida. No início de cada dia de trabalho, cada agente recebe uma lista de contratos atrasados.” (Agier e Assunção 2009)

Foi possível ao autor da monografia acompanhar um agente de crédito na Rocinha desde o momento em que este recebeu a lista dos clientes que estavam atrasados até seu retorno ao escritório. Reforçando a tese que o conhecimento local é relevante para o desempenho do contrato, o agente mostrou muita facilidade em lidar com os clientes que moram em localidades próximas a sua residência, até por conhecê-los pessoalmente, porém alguma incerteza sobre a situação dos clientes mais distantes.

A parte variável do salário de um agente depende do desempenho dos empréstimos que concedeu, sendo medida pelo fato de amortizações estarem atrasadas em mais de 30 dias. Conseqüentemente, os agentes dedicam muito tempo e esforço para garantir que seus empréstimos estejam em dia. No caso de um empréstimo continuar atrasado após diversas visitas do analista, se o atraso passar de 180 dias (passo 6) o montante é tratado como perda na contabilidade do VivaCred.

Portanto, percebe-se que o agente representa elemento essencial na conexão entre clientes e instituição nas duas fases de um empréstimo. A relação quase pessoal entre os analistas e seus respectivos clientes é responsável por mitigar grande parte das informações assimétricas encontradas no ciclo de crédito, especialmente se tratando de microcrédito.

4) Base de dados

Os dados utilizados neste trabalho são provenientes da base de dados oficial do VivaCred e contém dados sobre os clientes, analistas, contratos e negócios, de 1997 até 2007, para todas as seis agências. Estes podem ser divididos em duas bases de dados iniciais. A primeira delas - “Base de Endereços” - contém informações geográficas sobre a residência dos 40 analistas e de cada uma das mais de 35.000 pessoas que requereram um empréstimo, sendo esse aprovado ou não. Cada observação nesta base é referente a uma pessoa. A segunda - “Base de Contratos” - contém características dos mais de 31.000 contratos de crédito, clientes, analistas e negócios para onde o dinheiro foi direcionado. Cada observação nesta base é referente a um empréstimo.

Como o objetivo dessa pesquisa é entender a relevância da distância entre as residências do analista e do cliente para o desempenho do empréstimo, inicialmente foi necessário transformar os endereços destes em coordenadas geográficas – latitude e longitude. Sabendo quem são o cliente e o analista de cada contrato, e as coordenadas de suas respectivas residências, foi então possível calcular a distância entre estes dois pontos. Assim, no final deste processo, a base de dados (“Base de Contratos e Distâncias”) contava com variáveis de desempenho do empréstimo, variáveis de distância entre residências e variáveis de controle (características do cliente, analista, contrato e negócio) para cada um dos contratos. As bases serão analisadas individualmente a seguir.

A) Base de Endereços

A Base de Endereços do VivaCred é composta de mais de 35.000 observações, contendo as seguintes informações para cada pessoa:

Endereço – Subbairro¹³ – Bairro – Cidade – CEP

A fim de transformar esse conjunto de dados em coordenadas geográficas, foi utilizado um processo de geocodificação. Para isto, foi utilizada como ferramenta uma planilha do software Excel contendo um macro¹⁴ (código de programação) adaptado

¹³ Subbairro é o nome que será usado para localidades dentro de um mesmo bairro. Apesar de incomum para bairros oficialmente reconhecidos pela prefeitura, este é um conceito muito comum em áreas favelizadas.

¹⁴ Este macro pode ser encontrado em:

<http://www.blog.methodsinexcel.co.uk/2007/11/12/geocoding-in-excel-using-google-maps-api/>

para as necessidades dessa pesquisa, que processa as informações que são fornecidas a respeito de um endereço através do servidor do Google Maps e as transforma em coordenadas, fornecendo ainda o nível de precisão encontrado para aquele endereço. Estes níveis são categorizados por números de 1 a 8, seguindo a ordem de menor para maior precisão: País, Cidade, CEP, Rua, Endereço. Exceto no caso em que o endereço exato é achado pelo sistema, a coordenada fornecida é a mais precisa possível, sendo representada pelo ponto central da localidade encontrada

Entretanto, no formato original, grande parte da base apresentava informações que não seriam compreendidas pelo servidor. Como o VivaCred é uma instituição que realiza empréstimos para pessoas de baixa renda, muitos desses clientes são residentes de favelas, comunidades que não são reconhecidas oficialmente, portanto não constam na base de endereços do Google. Além disso, muitos endereços continham erros de digitação ou não estavam formatados corretamente.

Exemplos destes de erros podem ser:

- BECO 1 - 5 - 65 ## E
- ISABEL DOMINGUES SN ## QD.#A CS 10
- RUA 42(VILA CARIOCA)
- ETRADA DA GÁVEA ## CASA # FDS
- TRAV.OLIMPICA N°## CACHOPA
- RUA TRÊS CASA ## BECO ##
- TRAV. GLORIA ## AP.## BL ##

Assim, para passar estes endereços pelo macro, foi necessário primeiro realizar um intenso trabalho de limpeza e padronização de dados. Com isto, estava-se tentando atingir o modelo de endereço reconhecido pelo Google Maps. Como exemplo, o endereço da Pontifícia Universidade Católica do Rio de Janeiro, na formatação ideal para a planilha, seria:

Rua Marquês de São Vicente, 225 – Gávea – Rio de Janeiro – RJ – 22451-900

A maneira encontrada para realizar esta limpeza, formatação e padronização foi escrever códigos de programação no software Stata que realizam automaticamente os mesmos processos para toda a base de dados. Devido ao tamanho da base de dados, seria impossível analisar cada endereço, portanto os únicos dados tratados individualmente com relação ao endereço foram os dos analistas, devido a seu número reduzido (40) e conseqüente relevância para os cálculos. Cada analista participa de milhares de empréstimos, enquanto cada cliente participa de, em média, 3 contratos.

Com a base de dados já formatada nos moldes que maximizam o entendimento pelo servidor, todos os endereços foram submetidos à conversão para coordenadas. Assim, foram geradas tabelas que continham o endereço, subbairro, bairro, cidade, Estado, CEP e País do cliente, assim como suas coordenadas geográficas e seu nível de precisão. A base de dados contendo estas informações foi chamada de “BASE 2”. Vemos um exemplo na Tabela 2 a seguir:

Tabela 2: Endereços, Coordenadas e Nível de Precisão para a BASE 2

ID	Endereço	Subbairro	Bairro	Cidade	Estado	CEP	País	Precisão	Latitude	longitude2	Nível de Precisão
1	RUA PASSAGEM, ##		Rio das Pedras	RIO DE JANEIRO	RJ	22753720	BRAZIL	0			Não localizado
2			Rio de Janeiro	RIO DE JANEIRO	RJ		BRAZIL	4	-22.903539	-43.209587	Cidade
3	RUA ALTA ELVIRA TEIXEIRA, ##		Jacarepaguá	RIO DE JANEIRO	RJ	22713000	BRAZIL	0			Não localizado
4	RUA MONTE APRAZIVEL, ##		Tomas Coelho	RIO DE JANEIRO	RJ	21381170	BRAZIL	8	-22.873645	-43.306235	Endereço
5	ESTRADA DA DA GAVEA RUA DOIS, #		rocinha	RIO DE JANEIRO	RJ	22451261	BRAZIL	6	-22.990521	-43.252262	Rua
6	RUA SSA SENHORA DA PENHA, ##		Rio de Janeiro	RIO DE JANEIRO	RJ	20931000	BRAZIL	0			Não localizado
7	RUA BENTO LISBOA, ##			RIO DE JANEIRO	RJ	22221001	BRAZIL	8	-22.927144	-43.179353	Endereço
8	RUA NESTOR BECO OLIVEIRA, #			RIO DE JANEIRO	RJ	23515680	BRAZIL	0			Não localizado
9	RUA DA LADEIRA, ###	Cachopa	Rocinha	RIO DE JANEIRO	RJ	22451261	BRAZIL	0			Não localizado
10	RUA NARCISO LOTE, #			CAXIAS		22255250	BRAZIL	8	-22.506105	-43.203741	Endereço
11	TRAVESSA LIBERDADE, ##		rocinha	RIO DE JANEIRO	RJ	22451261	BRAZIL	6	-22.991336	-43.251942	Rua
12	RUA TRINTA E CINCO, ###			RIO DE JANEIRO	RJ	23595000	BRAZIL	8	-22.935781	-43.564285	Endereço

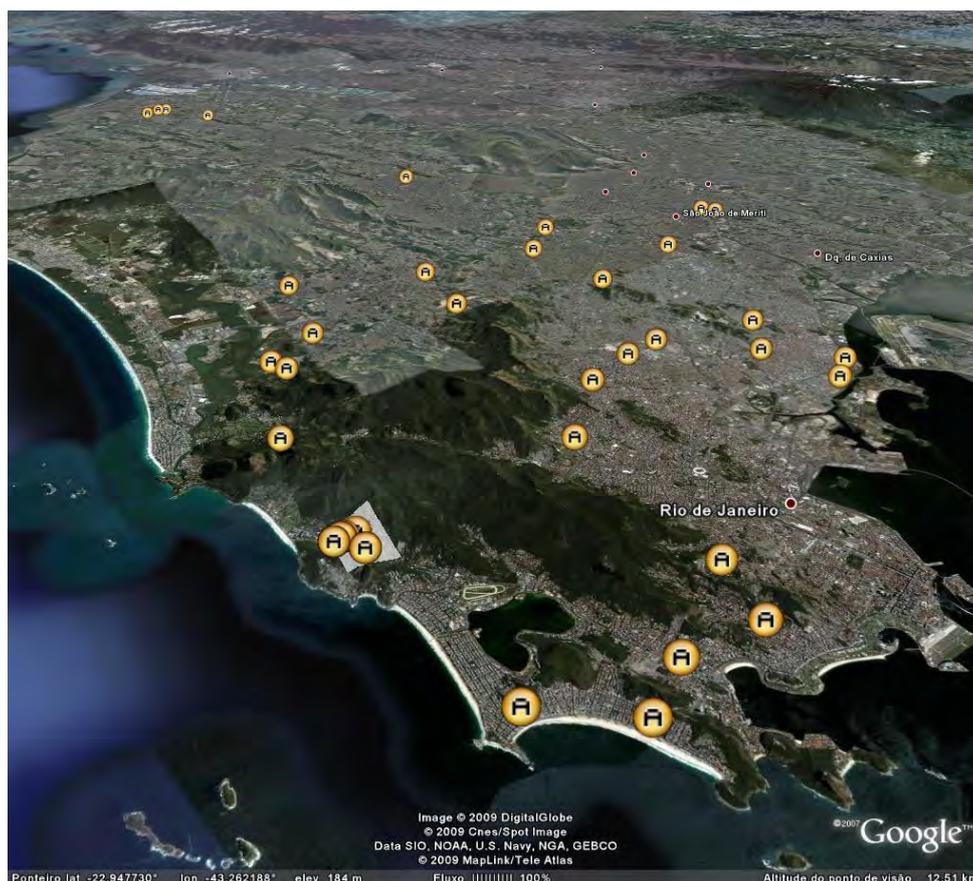
Pode-se observar que, mesmo depois de todo o trabalho de limpeza e padronização, o servidor não consegue localizar, ou localiza parcialmente, uma porção considerável dos endereços que lhe são fornecidos. Como pode ser visto na Tabela 3 abaixo, foram reconhecidos 16.210 dentre os 35.194 endereços submetidos à análise, o que representa 46,10% da amostra.

Tabela 3: Endereços retornados por categoria para BASE 2

Categoria	Nível de Precisão	Numero de observações	% do total
0	Não localizado	18955	53.90%
1	País	87	0.25%
2	Região	1	0.00%
3	Sub-Região	0	0.00%
4	Cidade	2079	5.91%
5	CEP	37	0.11%
6	Rua	4783	13.60%
7	Interseção	0	0.00%
8	Endereço	9223	26.23%
TOTAL		35.194	100.00%

Devido ao baixo percentual de endereços encontrados, foi realizada nova análise para entender o que havia de diferente entre os endereços que estavam sendo reconhecidos e os que não estavam. Primeiramente os dados foram observados de forma individual, a fim de se reconhecer padrões nos dados que não estavam sendo encontrados, e logo foi entendido que a maioria destes se localizava em favelas. Para isso, foi usado um site que, a partir do fornecimento de um endereço, mostra como o servidor lê os dados e verifica as possibilidades de coordenadas para este¹⁵. Posteriormente, para confirmar a análise inicial, foi utilizado como ferramenta de análise geográfica o Google Earth. Isto foi feito plotando-se pontos na superfície do globo terrestre através de suas coordenadas geográficas¹⁶, o que pode ser visto a seguir nas Figuras 3 e 4. As residências dos analistas estão definidas por pontos amarelos com a letra “A”, e as dos clientes por pontos azuis.

Figura 3: Localização dos agentes de crédito



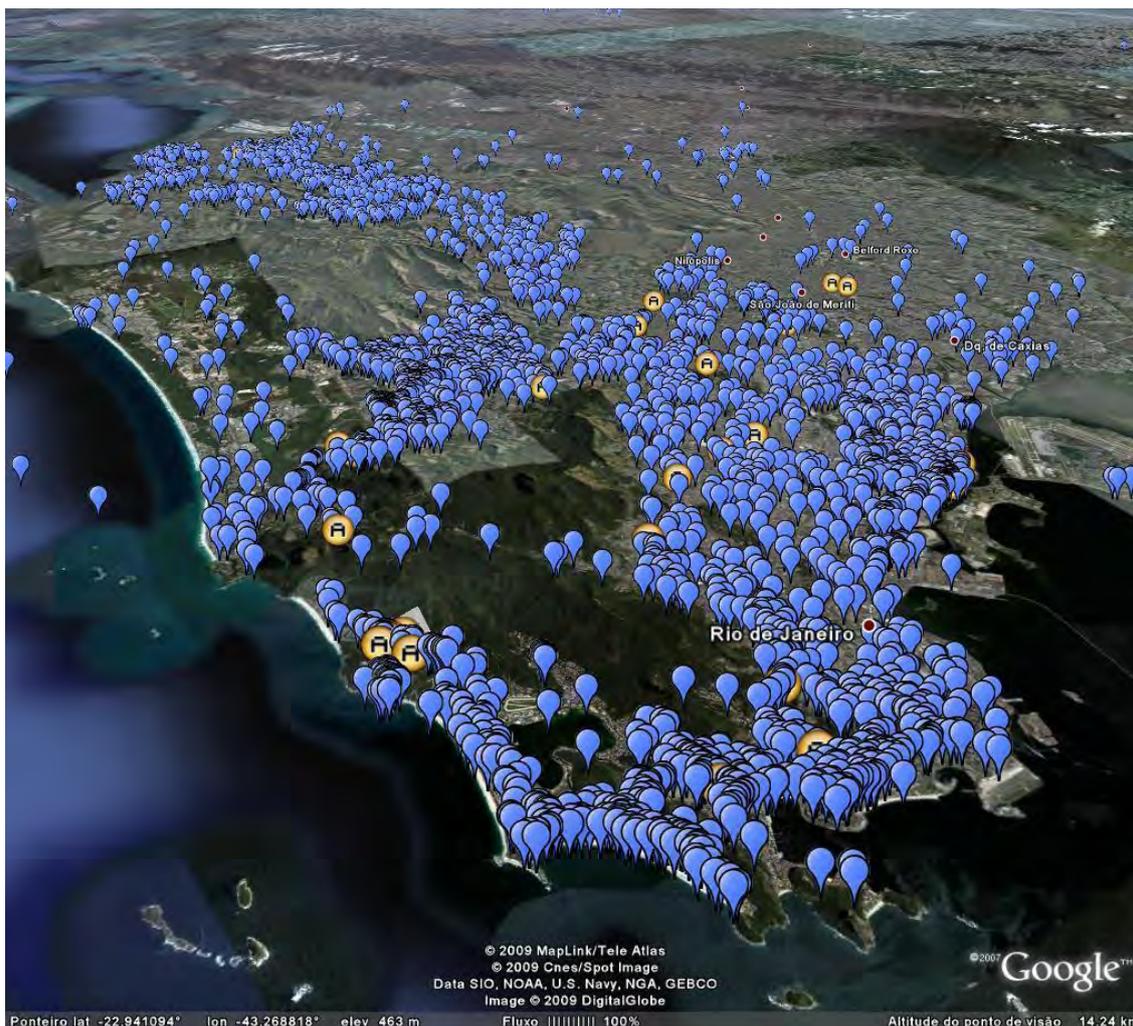
Fonte: Google Earth®

¹⁵ http://koti.mbnet.fi/ojalesa/geocode/prettygeocode_v2.htm

¹⁶ Agradeço a autorização para uso acadêmico do site:

<http://www.earthpoint.us/ExcelToKml.aspx>, assim como o suporte oferecido por seu criador, Bill Clark

Figura 4: Localização dos clientes e agentes de crédito



Fonte: Google Earth®

Com esta ferramenta foi possível navegar em três dimensões dentro do que antes era somente uma base de dados alfanumérica, permitindo a análise visual dos endereços encontrados. Isto confirmou que não estavam sendo localizadas principalmente as observações referentes às áreas de atuação das agências Rocinha, Rio das Pedras e Maré: por serem regiões onde as ruas, travessas e becos não são cadastrados pelos órgãos municipais, independentemente da limpeza dos endereços, tais logradouros não existem para o Google Maps.

Assim, por conter as coordenadas de grande parte dos endereços oficialmente reconhecidos e não conter as dos endereços de favelas – que representam em torno de metade do total de endereços - esta BASE 2 não poderia ser aceita como amostra satisfatória dos dados como um todo. Por existir diferença significativa entre moradias de favela ou não, caso a base fosse aceita haveria um viés de seleção, levando a serem

encontrados, nas regressões, resultados refletindo o efeito das distâncias nos empréstimos para somente um dos tipos de cliente.

******Para aumentar a quantidade e diversidade de dados localizados foi preciso encontrar outra maneira de se localizar estes endereços, e para isso, foi necessário entender qual era a precisão desejada. Como neste trabalho se está buscando entender a importância do conhecimento local do agente de crédito, e usando a distância entre duas residências como Proxy¹⁷ para isso, não se faz necessária uma precisão métrica: o relevante para entender o relacionamento entre o agente e seus clientes é saber quanto perto ou longe estes moram entre si¹⁸, de maneira categórica.

A solução encontrada, dada a restrição de informações contidas na base de dados, foi utilizar os subbairros para definir onde moram estas pessoas que não tiveram seus endereços reconhecidos automaticamente pelo servidor do Google. Favelas são geralmente divididas em localidades com nomes próprios, e com áreas razoavelmente bem definidas, que podem ser chamadas de subbairros. Na Rocinha, Maré e Rio das Pedras existem em torno de 20, 16 e 6 subbairros, respectivamente. Na base de dados, clientes que moram nestas áreas geralmente têm seus subbairros definidos no campo destinado a isso, funcionando como uma referência para localização da residência destes.

A fim de transformar a informação do subbairro dos clientes em coordenadas foi necessário entender a delimitação geográfica das áreas destes. Para isto foram utilizados mapas do Sistema de Assentamentos de Baixa Renda (SABREN)¹⁹, do Instituto Pereira Passos²⁰ e Wikimapas²¹. De posse dos mapas, foram consultados os agentes²² que participaram da definição destes subbairros no sistema do VivaCred, para que estes indicassem a área de cada um. Com a informação do perímetro destas localidades foram

¹⁷ Proxy é o nome dado em econometria para uma variável que serve como aproximação de outra. Por exemplo, consumo pode ser uma Proxy de riqueza.

¹⁸ O autor pôde observar o conceito de proximidade ao visitar, junto de um agente de crédito, a Favela da Rocinha, e observar que este conhecia pessoalmente todos os seus vizinhos mas a medida que se afastava de sua casa, menos gente parava para o cumprimentar. Este também demonstrava menos conhecimento sobre os clientes mais distantes, ainda que dentro da mesma comunidade.

¹⁹ <http://portalgeo.rio.rj.gov.br/sabren/index.htm>

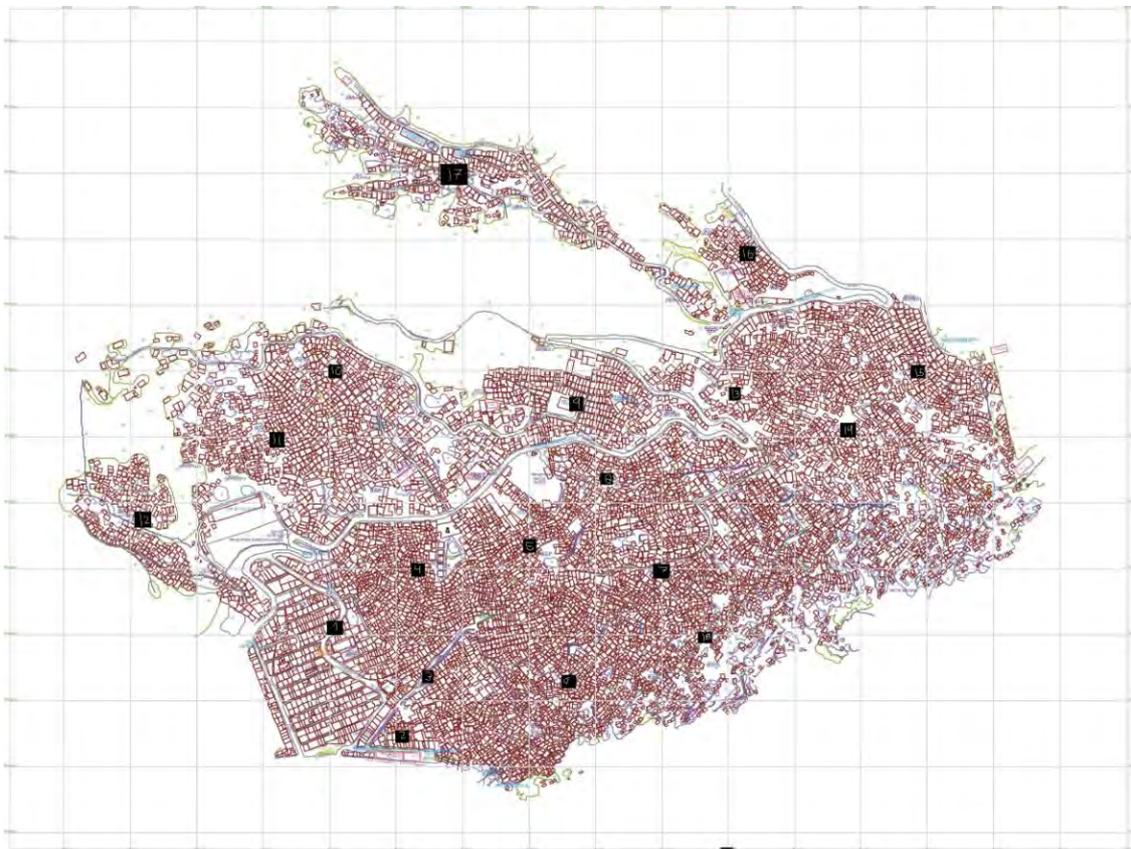
²⁰ <http://www.rio.rj.gov.br/ipp/>

²¹ <http://wikimapia.org/#lat=-22.8619796&lon=-43.2374239&z=15&l=9&m=b>
(para a Favela da Maré)

²² Antônio César Martins, da agência Rocinha, e “Tavares” da agência Rio das Pedras.

encontrados centróides²³ para elas. Pode-se observar na Figura 5 o mapa da Rocinha já com os centróides de subbairros marcados por quadrados pretos.

Figura 5: Mapa da Rocinha com centróides de subbairros definidos.



Fonte: SABREN / Instituto Pereira Passos.

Este mapa foi posteriormente sobreposto ao globo terrestre no Google Earth, como se observa na Figura 6, e as coordenadas de cada centróide anotadas. Processos similares foram realizados para todos os subbairros que aparecem com mais frequência nos cadastros das agências Rocinha, Maré e Rio das Pedras, e assim criada uma tabela de subbairros e suas respectivas coordenadas. Com isto, foi criada uma nova base de dados, chamada de BASE 3, aonde foi imputada a cada morador destas localidades a coordenada de seu subbairro, assim como um nível de precisão para esta. Nas áreas em que a região é extensa – como Cidade de Deus ou Complexo do Alemão – foi escolhido o nível 7, e para áreas específicas – como as sub-regiões da Rocinha – foi escolhido o nível de precisão 9. O motivo desta escolha vem da comparação com a BASE 2. Acredita-se que a precisão das coordenadas em grandes subbairros seja menor do que a

²³ Coordenada geográfica central de uma área.

dos endereços encontrados pelo macro. Entretanto, para o objetivo desta pesquisa, um subbairro bem definido pode ser considerado mais relevante do que o endereço, dado que o primeiro foi verificado individualmente e o segundo através de automações.**

Nesta BASE 3 foram relacionadas coordenadas para 11.050 pessoas, sendo 1.323 em nível 7 e 9.727 em nível 9, como pode ser visto na tabela 4 a seguir.

Figura 6: Mapa da Rocinha sobreposto em sua devida localização no globo terrestre.



Fonte: Google Earth

Tabela 4: Endereços retornados por categoria para BASE 3

Categoria	Nível de Precisão	Numero de observações	% do total
0	Não localizado	24.144	68.60%
7	Grande Subbairro	1.323	3.76%
9	Pequeno Subbairro	9.727	27.64%
TOTAL		35.194	100.00%

No final deste processo constavam duas bases de dados bastante diferentes. A primeira delas, BASE 2, continha coordenadas para grande parte dos endereços em bairros oficialmente reconhecidos pela prefeitura, porém quase nenhuma para endereços em favelas. A segunda, BASE 3, continha grande parte das coordenadas de favelas, porém nenhuma de endereços oficiais. Deste modo, as bases tinham vieses de seleção inversos, tornando possível e necessário juntar as duas em uma só base, que foi chamada de BASE 23.

A BASE 23 foi construída realizando uma fusão das duas bases antecedentes. Foram aproveitados os dados com maior qualidade de cada uma das anteriores, ou seja, foram comparadas as bases e escolhidas as coordenadas que haviam sido encontradas com maior nível de precisão. Assim, obteve-se uma base com número expressivamente maior de observações e com menor viés de seleção. A Tabela 5 abaixo resume as bases anteriores e mostra o resultado final para a BASE 23.

Tabela 5: Endereços retornados por categoria para todas as bases.

Categoria	Nível de Precisão	Numero de observações da BASE 2	Numero de observações da BASE 3	Numero de observações da BASE 23	BASE 23 % do total
0	Não localizado	18955	24.144	10.416	29.60%
1	País	87	0	87	0.25%
2	Região	1	0	1	0.00%
3	Sub-Região	0	0	0	0.00%
4	Cidade	2079	0	2048	5.82%
5	CEP	37	0	36	0.10%
6	Rua	4783	0	2937	8.35%
7	Subbairro Grande	0	1.323	1083	3.07%
8	Endereço	9223	0	8859	25.17%
9	Subbairro Pequeno.	0	9727	9727	27.64%
TOTAL		35194	35194	35194	100.00%

Percebe-se que pequena parte dos dados que haviam sido encontrados na primeira foram encontrados na segunda, e vice-versa, demonstrando a complementaridade das bases. O número de endereços não encontrados passou de aproximadamente 50% e 70%, na primeira e segunda bases respectivamente, para em torno de 30% na terceira. Entretanto, se os dados para os quais foram encontradas somente coordenadas centrais do país ou cidade fossem utilizados nos cálculos, estes estariam distorcendo nossa amostra de distâncias. Logo, só foram mantidos os dados com qualidade igual ou superior a 5, ou seja, quando pelo menos foi encontrado o CEP do indivíduo.

Para analisar a dispersão dos dados que foram encontrados, mais uma vez foi utilizado o Google Earth. Nesse momento se fez relevante dividir as observações por agência para entender a área de abrangência de cada uma, e perceber se as coordenadas encontradas para as pessoas se dispunham ao redor das agência que as cadastraram. Para cada agência foi definida uma cor seguindo a Tabela 6. O resultado pode ser visto nas Figuras 7, 8 e 9 abaixo.

Tabela 6: Legenda das cores das agências

Rocinha	
Glória	
Maré	
Rio das Pedras	
Santa Cruz	
Macaé	

Figura 7: Área metropolitana do Rio de Janeiro, Zona sul à frente e Oeste ao fundo.

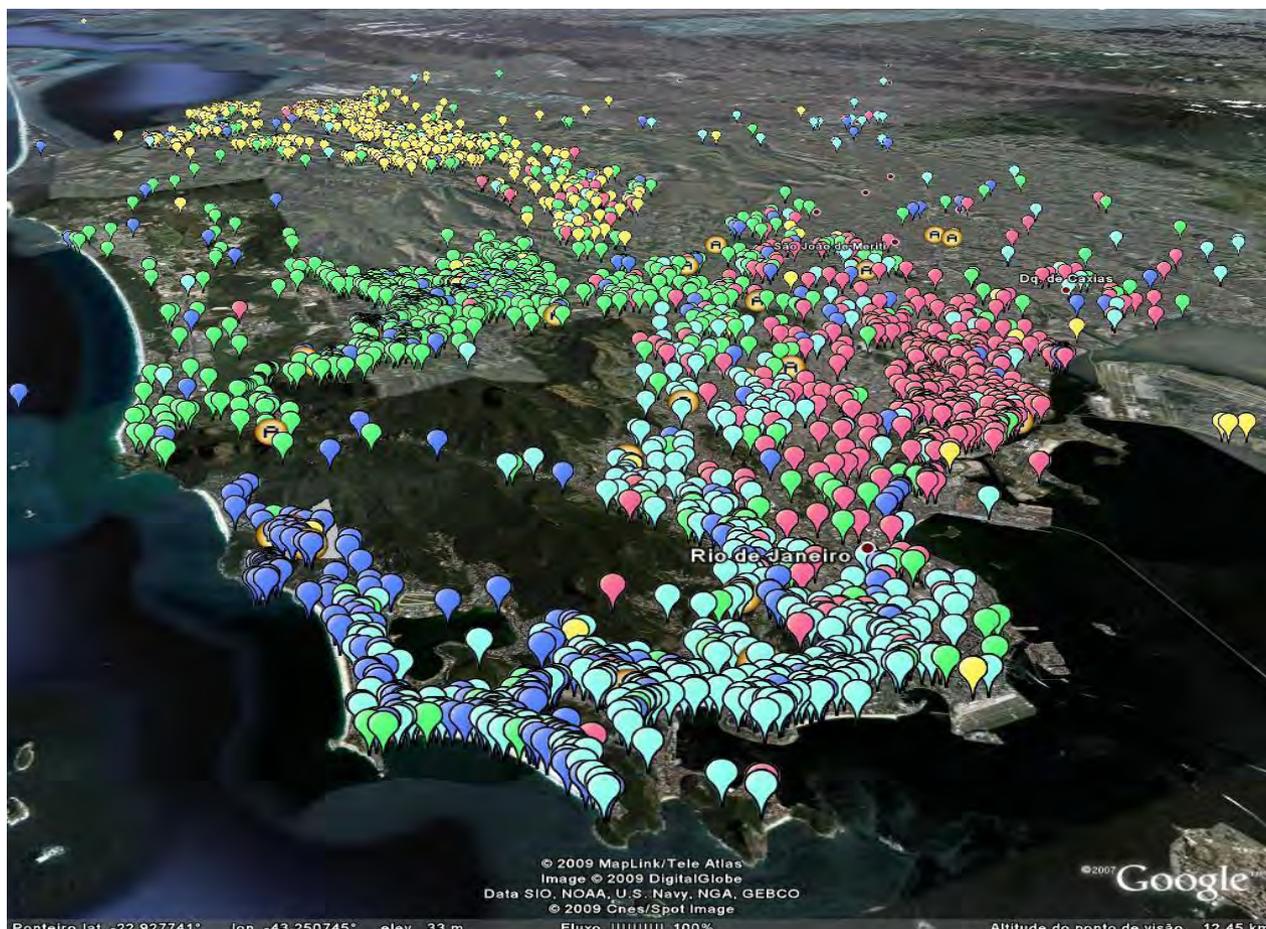
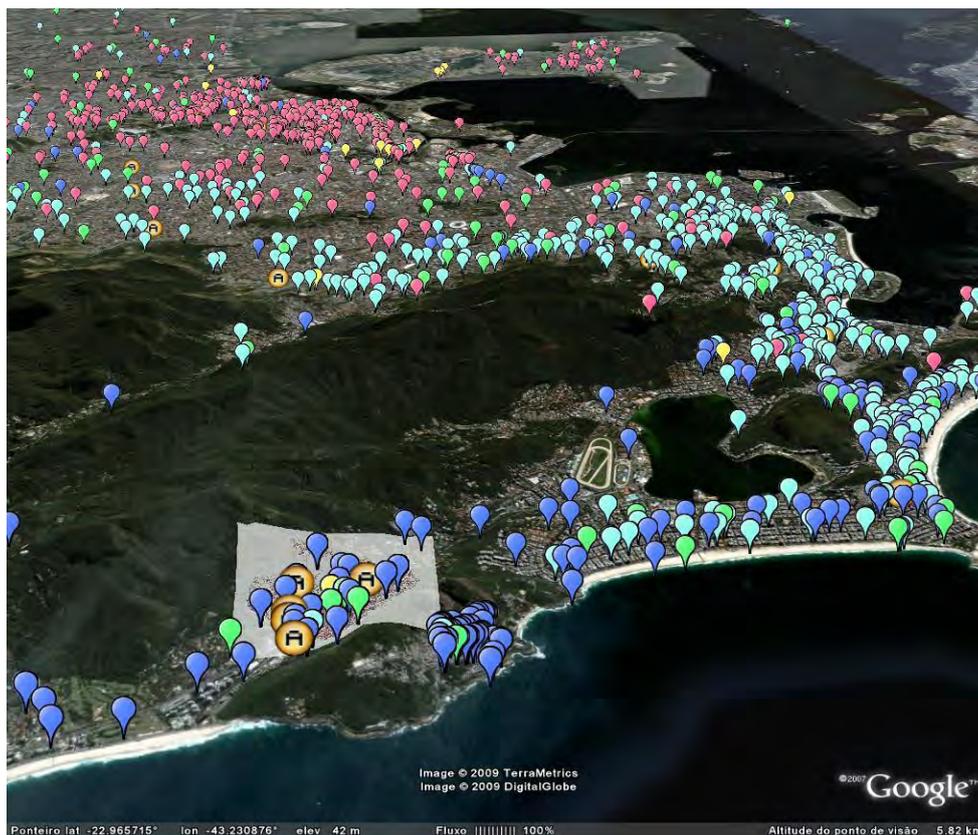


Figura 8: Rio de Janeiro, Zona Sul à frente, Zona Norte ao fundo.



Fonte: Google Earth®/ TerraMetrics, Digital Globe

Figura 9: Rio de Janeiro, Favela da Maré (Zona Norte) à frente, Zona Oeste ao fundo.



Fonte: Google Earth®/ Data Sio, NOAA, U.S. Navy, NGA, GEBCO

As figuras acima sugerem que há precisão considerável nos dados encontrados para a BASE 23. Isto comprova que através dos métodos de localização de coordenadas utilizados foi possível melhorar significativamente a qualidade e quantidade de informações. Assim, ao final do processo de transformação de endereços em coordenadas, foram geradas três bases diferentes, sendo a terceira uma fusão das duas primeiras. A tabela 7 resume as características destas.

Tabela 7: Características das bases de coordenadas.

BASE	Endereços não localizados	Endereços com precisão ≥ 5	% Endereços com precisão ≥ 5	Método de criação	Viés de Seleção (análise geográfica)
BASE 2	18.955	14.043	46,14%	Endereços passados pelo servidor do Google Maps após limpeza, formatação e padronização.	Maioria de endereços de localidades oficiais, poucos de favelas
BASE 3	24.144	11.050	31,40%	Coordenadas centróides dos subbairros imputadas manualmente.	Maioria de endereços de favelas, nenhum dos oficiais
BASE 23	10.416	22.642	64,33%	Fusão da BASE 2 e BASE 3, mantendo a coordenada encontrada com maior precisão.	Balço entre endereços oficiais e de favelas.

B) Base de Contratos

A Base de Contratos do VivaCred é composta de mais de 31.000 observações, cada uma sendo um empréstimo. Estes foram concedidos por 40 analistas a 11.620 clientes durante os 11 anos de operações para os quais se tem informações. Cada observação é composta de dados que podem ser divididos em quatro grupos de características: cliente, analista, contrato e negócio.

É nesta base que se encontra a Proxy para desempenho do contrato, a ser usada como variável dependente nas regressões. Esta é uma variável *dummy* para o caso de alguma parcela de amortização ter atrasado em mais de 30 dias. O motivo da escolha desta informação é o fato de 30 dias serem o prazo tolerado para atraso de uma parcela. O salário dos analistas é dividido em uma parte fixa e outra variável. Caso um cliente atrase mais do que 30 dias, a parte variável do salário será afetada negativamente. Além disso, este é o prazo considerado máximo a ser afetado por ações do analista.

Empréstimos atrasados mais de 180 dias são considerados como perda para a instituição, e poderiam ser usados como Proxy para desempenho, mas não acredita-se que as escolhas de repagar uma parcela de um empréstimo no prazo de 6 meses sejam afetadas pelos agentes e sim por fatores externos.

Também constam da Base de Contratos as informações usadas como variáveis de controle nas regressões. Como está se buscando mensurar a influência da distância entre residências do analista e cliente no desempenho do contrato de crédito, é preciso tentar eliminar deste coeficiente a influência de outros fatores. Por exemplo, podemos rodar uma regressão para entender o efeito da idade da pessoa em sua renda, porém, como existe correlação entre educação e idade, se não controlarmos essa regressão para educação, podemos estar afirmando que o efeito da idade é maior do que de fato é. Para esta pesquisa, busca-se eliminar o efeito de outras variáveis relevantes para o desempenho do contrato de dentro do coeficiente de variável distância. Detalha-se a seguir os quatro grupos de informação individualmente.

O primeiro grupo de informações, sobre o cliente que está tomando o empréstimo, contém informações pessoais como o sexo deste, se é casado, se tem dependentes e sua idade. Além destas, também contém informações financeiras como sua renda extra (não proveniente de sua atividade principal) corrigida pelo Índice de Preços ao Consumidor (IPC) para Reais de 1997, o consumo mensal de sua família, também corrigido pelo IPC para Reais de 1997, e se este tem conta corrente em banco.

Sobre o analista, o grupo de informações mais completo dentre os quatro, constam as seguintes informações: o sexo, se é casado, se tem dependentes, seu nível de escolaridade, sua idade, e se mora em área favelizada. Esta última informação é relevante para saber se o analista tem conhecimento da realidade cotidiana de um cliente que mora em favela. Muitas vezes a lógica de negócios dentro de favelas pode ser diferente do que se está acostumado fora delas, por ser mais informal. Além destas também constam informações sobre a vida profissional pregressa do agente. Sabe-se quantos anos este tem de experiência profissional geral, se trabalhava na área administrativa do VivaCred antes de se tornar analista e quantos anos tem de experiência em vendas.

Sobre o negócio para onde está sendo destinado o dinheiro está indicado se este é oficializado - negócios oficiais têm Cadastro Nacional de Pessoa Jurídica (CNPJ) – o número de empregados e seu balanço, corrigido pelo IPC para Reais de 1997.

Sobre os contratos, é importante ressaltar que entre eles existe grande semelhança. Todos os contratos efetuados pelo VivaCred que utilizamos nesta pesquisa são concedidos a indivíduos, e não a grupos, como é muito comum se tratando de microcrédito. Além disso, de todos os clientes é cobrada a mesma taxa de juros de 3.9% ao mês, mais uma taxa de inscrição, que varia de 5% para novos clientes a 3% para clientes antigos e com boa reputação. Os termos que diferem de um contrato para outro, e sobre os quais temos informação, são: se teve fiador (garantidor), o número de parcelas de amortização, se este foi usada para investimentos em capital fixo, se foi usado para repagamento de empréstimos e seu valor, corrigido pelo IPC para Reais de 1997.

C) Base de Contratos e Distâncias

Para poder rodar as regressões de interesse para este estudo, foi necessário integrar as bases acima descritas de maneira a se obter no final deste processo uma única base consolidada, chamada de Base de Contratos e Distâncias. Esta contém a variável dependente – a *dummy* de atraso maior do que 30 dias – e as coordenadas a serem usadas para se obter as distâncias entre residências, que é a principal variável independente. Além disso, também estão incluídas todas as variáveis de controle contendo informações sobre clientes, analistas, contratos e negócios.

Sabendo quem foram o analista e o cliente de um determinado contrato, e tendo as coordenadas de suas residências, é possível calcular a distância entre esses dois pontos. Este cálculo é feito através de um método chamado de “vôo do pássaro” ou “grande círculo”, em que se mede a distância entre dois pontos não levando em consideração o relevo terrestre, como o vôo de um pássaro. É calculada a menor distância entre dois pontos levando em consideração a curvatura da Terra. A fórmula usada para isso é:

$$\text{Distância} = 6371.1 * ((2 * \text{ASIN}(\text{SQRT}(\text{SIN}((\text{RADIANOS}(\text{LAT1}) - \text{RADIANOS}(\text{LAT2}))/2)^2 + \text{COS}(\text{RADIANOS}(\text{LAT1})) * \text{COS}(\text{RADIANOS}(\text{LAT2})) * (\text{SIN}((\text{RADIANOS}(\text{LON1}) - \text{RADIANOS}(\text{LON2}))/2)^2))))))$$

Nesta fórmula, 6371.1 que multiplica todo o resto é o raio da Terra em Kilômetros, fazendo com que a distância resultante do cálculo também esteja em Km. É importante notar que a fórmula utiliza latitudes e longitudes em Radianos. Para realizar

esta conversão, dado que temos coordenadas em padrão decimal, devemos utilizar a seguinte fórmula:

$$\text{CoordenadaRadianos} = \text{CoordenadaDecimal} * \frac{\Pi}{180}$$

Entretanto, como visto anteriormente, apesar de terem sido encontradas todas as coordenadas das residências de analistas, o mesmo não é verdade para as coordenadas das residências dos clientes. Para estes, as possibilidades são:

1. Não ter se encontrado coordenada alguma.
2. Ter se encontrado coordenada de precisão baixa, que foi descartada.
3. Ter se encontrado coordenada de precisão satisfatória somente no processo de criação da BASE 2
4. Ter se encontrado coordenada de precisão satisfatória somente no processo de criação da BASE 3
5. Ter se encontrado coordenada de precisão satisfatória tanto no processo de criação da BASE 2 quanto no processo de criação da BASE 3.

Nos casos 1 e 2, por não existirem coordenadas ou estas serem insatisfatórias, não foi possível calcular distâncias entre cliente e analista. Nos casos 3 e 4, as distâncias foram calculadas entre a coordenada da residência do analista e a coordenada encontrada na respectiva base. Para o caso 5, foram calculadas as distâncias tanto para as coordenadas encontradas na BASE 2 quanto na BASE 3. Nos casos 3, 4 e 5 também foi calculada a distância para a BASE 23.

Dessa maneira, ao final de toda consolidação e cálculos, cada observação contava com a variável explicitando se houve atraso em alguma das parcelas da amortização, as distâncias (para os casos em que estas foram calculadas) e as variáveis de controle. A tabela 8 resume todas as variáveis, separadas por grupo de características, e seus tipos e valores possíveis.

Para se analisar as bases de coordenadas, e saber se as médias referentes às observações para as quais foram encontradas coordenadas na BASE 23 são similares as da base como um todo, foi construída a Tabela 9, de médias e diferenças entre médias das bases, incluindo também a significância da diferença entre elas. Nesta, pode-se

confirmar os vieses de seleção inferidos através da análise geográfica para a BASE 2 e BASE 3. Pode-se notar também que a BASE 23 tem menos viés de seleção que as duas bases que a antecederam.

Tabela 8: Descrição das variáveis da Base de Contratos e Distâncias.

Variável	Tipo	Valores Possíveis	Descrição
Variáveis de desempenho do empréstimo			
Atraso > 30 dias	Dummy	1 - Sim ; 0 - Não	Alguma das parcelas do empréstimo atrasou mais de 30 dias.
Variáveis de distância			
Distância 2 (Km)	Contínua	0 - 254	Distância da residência do analista à do cliente, usando as coordenadas disponíveis na BASE 2, em Km.
Distância 3 (Km)	Contínua	0 - 46	Distância da residência do analista à do cliente, usando as coordenadas disponíveis na BASE 3, em Km.
Distância 23 (Km)	Contínua	0 - 254	Distância da residência do analista à do cliente, usando as coordenadas disponíveis na BASE 23, em Km.
Características do Analista			
Homem	Dummy	1 - Sim ; 0 - Não	Analista é homem.
Casado	Dummy	1 - Sim ; 0 - Não	Analista é casado.
Dependentes	Dummy	1 - Sim ; 0 - Não	Analista tem dependentes.
Mora em favela	Dummy	1 - Sim ; 0 - Não	Analista mora em área favelizada.
Nível Superior	Dummy	1 - Sim ; 0 - Não	Analista tem nível superior de escolaridade.
Experiência anterior (anos)	Contínua	0.5 - 27	Anos de experiência profissional anterior do analista.
Experiência administrativa	Dummy	1 - Sim ; 0 - Não	Analista trabalhava anteriormente na área administrativa da instituição.
Experiência em vendas	Contínua	0 - 11	Anos de experiência profissional anterior do analista em vendas.
Idade	Contínua	18 - 61	Idade do analista.
Características do Cliente			
Homem	Dummy	1 - Sim ; 0 - Não	Cliente é homem.
Casado	Dummy	1 - Sim ; 0 - Não	Cliente é casado.
Dependentes	Dummy	1 - Sim ; 0 - Não	Cliente tem dependentes.
Idade	Contínua	17 - 108	Idade do cliente.
Renda extra da família (R\$)	Contínua	0 - 17.364	Renda da família proveniente de outras fontes (Reais de 1997; corrigido pelo IPC).
Consumo da família (R\$)	Contínua	0 - 7.323	Consumo mensal da família (Reais de 1997; corrigido pelo IPC).
Tem conta corrente	Dummy	1 - Sim ; 0 - Não	Cliente tem conta corrente em banco.
Características do Negócio			
Negócio registrado	Dummy	1 - Sim ; 0 - Não	O negócio é registrado. Tem CNPJ.
Numero de empregados	Contínua	0 - 120	Número de empregados do negócio.
Balanço do Negócio	Contínua	0 - 52.255	Balanço do negócio (Reais de 1997; corrigido pelo IPC).
Características do Contrato de Crédito (Empréstimo)			
Valor do empréstimo	Contínua	53 - 17.780	Valor do empréstimo (Reais de 1997; corrigido pelo IPC)
Fiador	Dummy	1 - Sim ; 0 - Não	Empréstimo teve fiador (garantidor).
Número de parcelas	Contínua	1 - 77	Número de parcelas de amortização do empréstimo.
Usado para capital fixo	Dummy	1 - Sim ; 0 - Não	Empréstimo foi usado para investimento em capital fixo.
Repagamento de dívida	Dummy	1 - Sim ; 0 - Não	Empréstimo foi usado para repagamento de dívidas.

Tabela 9: Médias, diferenças e signific das variáveis da Base de Contratos e Distâncias.

Base de Dados	TODOS	Diferença Sign. BASE 2 (B2 / Todos) - da			Diferença Sign. BASE 3 (B3 / Todos) - da			Diferença Sign. BASE 23 (B23 / Todos) - da		
		1	difer.		1	difer.		1	difer.	
Observações	31691	11945			12765			22059		
Variáveis de desempenho do empréstimo										
Atraso > 30 dias	0.0861	0.1094	27.06%	***	0.0723	-16.03%	***	0.0902	4.76%	***
Variáveis de distância										
Distância 2 (Km)	12.4806	12.4806								
Distância 3 (Km)	6.1356				6.1356					
Distância 23 (Km)	9.4937							9.4937		
Características do Analista										
Homem	0.5260	0.4568	-13.15%	***	0.6517	23.89%	***	0.5397	2.60%	***
Casado	0.5457	0.5895	8.02%	***	0.4192	-23.18%	***	0.5120	-6.16%	***
Dependentes	0.5764	0.6198	7.54%	***	0.4888	-15.19%	***	0.5480	-4.93%	***
Mora em favela	0.6910	0.5898	-14.65%	***	0.8523	23.35%	***	0.7078	2.44%	***
Nível Superior	0.0725	0.0946	30.46%	***	0.0478	-34.10%	***	0.0722	-0.41%	***
Experiência anterior (anos)	7.9379	7.8999	-0.48%	***	7.7269	-2.66%	***	7.8511	-1.09%	***
Experiência administrativa	0.0904	0.0757	-16.32%	***	0.1162	28.46%	***	0.0910	0.66%	***
Experiência em vendas	3.0478	2.8330	-7.05%	***	3.3733	10.68%	***	3.0700	0.73%	**
Idade	31.287	31.016	-0.87%	***	30.758	-1.69%	***	30.991	-0.94%	***
Características do Cliente										
Homem	0.5040	0.4874	-3.29%	***	0.5197	3.12%	***	0.4992	-0.95%	***
Casado	0.4757	0.4407	-7.36%	***	0.4987	4.84%	***	0.4707	-1.04%	***
Dependentes	0.5220	0.4913	-5.88%	***	0.5958	14.14%	***	0.5367	2.81%	***
Idade	42.376	42.541	0.39%	*	42.367	-0.02%		42.568	0.45%	***
Renda extra da família (R\$)	217.71	224.75	3.24%	***	221.16	1.59%		217.83	0.06%	***
Consumo da família (R\$)	525.08	571.26	8.79%	***	460.52	-12.30%	***	512.71	-2.36%	***
Tem conta corrente	0.0674	0.0600	-10.98%	***	0.0722	7.00%	***	0.0661	-1.92%	***
Características do Negócio										
Negócio registrado	0.0580	0.0735	26.65%	***	0.0427	-26.44%	***	0.0574	-1.05%	***
Numero de empregados	0.6268	0.7281	16.17%	***	0.4321	-31.06%	***	0.5878	-6.21%	***
Balanço do Negócio	939.24	980.44	4.39%	***	803.38	-14.46%	***	888.50	-5.40%	***
Características do Contrato de Crédito (Empréstimo)										
Valor do empréstimo	1014.20	1034.14	1.97%	***	1024.31	1.00%		1006.40	-0.77%	**
Fiador	0.9304	0.9310	0.06%		0.9318	0.14%		0.9286	-0.20%	**
Número de parcelas	9.0376	8.8736	-1.82%	***	9.5788	5.99%	***	9.1955	1.75%	***
Usado para capital fixo	0.3201	0.3215	0.44%		0.3390	5.91%	***	0.3240	1.24%	**
Repagamento de dívida	0.0945	0.0928	-1.85%		0.0769	-18.60%	***	0.0876	-7.28%	***

* Significante a 10%, ** Significante a 5%, *** Significante a 1%.

A hipótese nula para o teste de hipóteses foi a diferença da media das observações encontradas e não encontradas na base ser igual a 0. Assim, quanto maior a significância, menor a semelhança.

Ao observar a base original, contendo todos os 31.691 contratos, nota-se que 8.6% dos empréstimos tiveram alguma parcela atrasada mais de 30 dias. Para os 11.945 contratos que tiveram distâncias identificadas através das coordenadas da BASE 2, 10.94% atrasaram alguma parcela. Já para os contratos que tiveram distâncias identificadas através de coordenadas da BASE 3, somente 7,23% dos empréstimos tiveram algum atraso. Ao observar as distâncias médias de cada base nota-se que a média para a BASE 2 é 12,48 Km e para a BASE 3, 6.13 Km. Com isto, pode-se inferir que exista relação entre distância e atraso, o que será analisado nas regressões a seguir.

No conto geral, percebe-se pela diferença de médias entre as bases e os contratos como um todo, que a BASE 23 gera uma aproximação para a base original bastante próxima em termos numéricos. Com estas diferenças nota-se também que as bases 2 e 3 de fato tem vieses opostos em relação à base completa.

Apesar disto, percebe-se que as diferenças entre médias são em geral significantes a 1%. Como está se usando como hipótese nula a diferença das médias ser igual a 0, esta é rejeitada na maioria dos casos. Isto quer dizer que não foi possível criar uma base onde os contratos com distâncias definidas têm médias estatisticamente iguais às médias da base original. Isto demonstra que a capacidade de identificar os endereços não provocou uma seleção aleatória da amostra. Se isso acontecesse, as médias não deveriam ser estatisticamente significativas. Dessa forma, os resultados mostram que ao realizar as regressões, é importante controlar para diferenças observáveis na amostra.

5) Resultados Empíricos:

A partir dos dados da Base de Contratos e Distâncias, foram rodadas regressões a fim de analisar a relevância da distância entre a residência do analista e a do cliente no desempenho do empréstimo, definido pela variável *dummy* de atraso maior do que 30 dias. A equação 1 mostra o modelo da regressão. A variável dependente foi regredida em uma constante e variáveis independentes, sendo que para cada uma das colunas das Tabelas 10 e 11, a regressão foi rodada com controles diferentes. Na coluna 1 foi rodada sem controles, nas colunas (2) – (6) foi rodada com controles específicos para cada grupo de informações, e na coluna (7) foi rodada com todos os controles. Como visto anteriormente, dado a falta de uma amostra aleatória que seja estatisticamente semelhante à base original, é importante adicionarmos todos os controles disponíveis, nos levando a prestar atenção especial ao coeficiente encontrado na coluna (7).

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i, \quad [1]$$

Tabela 10: Regressão de distância medida através de coordenadas da BASE 23 em *dummy* de atraso maior que 30 dias, com e sem controles.

Ln D23	Atraso > 30 dias	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Distância	0.01 (9.00)***	0.009 (7.98)***	0.007 (4.21)***	0.01 (8.99)***	0.01 (8.62)***	0.008 (6.14)***	0.003 (1.74)*
Características incluídas:	Cliente		■					■
	Analista			■				■
	Crédito				■			■
	Negócio					■		■
	Agência (Dummy)						■	■
Observações		22059	22059	18492	22059	22045	22059	18481
R2		0.00	0.02	0.03	0.01	0.01	0.00	0.05

Estadísticas T robustas dentro dos parênteses

* significante a 10%; ** significante a 5%; *** significante a 1%

A tabela 10 apresenta os resultados da estimação da equação [1] com as distâncias encontradas para coordenadas da BASE 23. A coluna (1), sem controles, sugere que se

dobrarmos a distância entre cliente e analista, o atraso aumenta em 1 ponto percentual. O coeficiente estimado não apenas é estatisticamente significativo, mas também economicamente relevante considerando que a média de atrasos de 30 dias é 8.6%. Entretanto, como a medida de distância entre analistas e clientes pode estar correlacionada com outras características observadas do cliente, do analista, do negócio ou mesmo da agência do Vivacred é importante adicionar controles. Dessa forma, as colunas (2) - (6) acrescentam variáveis observáveis como controles adicionais para a estimação da equação [1]. Em todos os casos, o coeficiente da distância não só se mantém estatisticamente significativo mas também se mantém estável, oscilando entre 0.007 e 0.009. Quando todos os controles são adicionados conjuntamente, na regressão de maior importância deste estudo, o coeficiente cai para 0.003, mantendo-se significativo.

Para realizar exercício de robustez, a mesma regressão foi rodada usando as observações para as quais foram encontradas distâncias na BASE 2.

Tabela 11: Regressão de distância medida através de coordenadas da BASE 2 em *dummy* de atraso maior que 30 dias, com e sem controles.

Ln D2	Atraso > 30 dias	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Distância	0.009 (4.34)***	0.009 (4.29)***	0.007 (2.77)***	0.009 (4.30)***	0.009 (4.22)***	0.008 (3.53)***	0.006 (2.09)**
Características incluídas:	Cliente		■					■
	Analista			■				■
	Crédito				■			■
	Negócio					■		■
	Agência (Dummy)						■	■
Observações		11945	11945	10360	11945	11940	11945	10356
R2		0.00	0.02	0.04	0.01	0.00	0.00	0.06

Estatísticas T robustas dentro dos parênteses

* significativa a 10%; ** significativa a 5%; *** significativa a 1%

A Tabela 11 apresenta os resultados da estimação da equação [1], desta vez com os dados da BASE 2. A coluna (1) sugere que se dobrarmos a distância entre cliente e analista, o atraso aumenta em 0.9 pontos percentuais. O coeficiente estimado, assim como nas regressões usando a BASE 23, não apenas é estatisticamente significativo, mas economicamente relevante considerando que a média de atrasos de 30 dias é 8.6%.

Entretanto, a medida de distância entre analistas e clientes pode estar correlacionada com outras características observadas do cliente, do analista, do negócio ou mesmo da agência do Vivacred. Dessa forma, as colunas (2)-(6) acrescentam variáveis observáveis como controles adicionais para a estimação da equação [1]. Em todos os casos, o coeficiente da distância não só se mantém estatisticamente significativo mas também se mantém estável, oscilando entre 0.007 e 0.009. Quando todos os controles são adicionados conjuntamente, o coeficiente cai para 0.006, mantendo-se significativo.

Estes resultados sugerem que estavam corretas as hipóteses iniciais de que existe de fato correlação entre a distância das residências e o desempenho do contrato, sendo o aumento da distância uma possível explicação para atraso do repagamento.

6) Conclusão:

Este trabalho apresentou evidência empírica de que a proximidade entre um analista de crédito e seus clientes tem efeito no desempenho de um empréstimo. A estratégia empírica adotada foi utilizar a distância entre as residências como principal variável independente, e regressi-la em uma variável *dummy* de atraso do empréstimo. Usando a base de dados de clientes e empréstimos do VivaCred, foi possível demonstrar que quanto mais perto está o agente de crédito de seu cliente, menor será a probabilidade desse determinado cliente atrasar alguma das parcelas da amortização de seu empréstimo em mais de 30 dias.

Mais do que isso, a distância entre residências pode ser entendida como Proxy para conhecimento local, e conhecimento local como Proxy para conhecimento específico. Dessa maneira, este estudo sugere que, para obter-se bom desempenho em contratos de crédito – pela perspectiva da firma - é relevante a participação de intermediários que tenham conhecimento específico sobre seus clientes.

7) Bibliografia:

- AGIER, I. E ASSUNÇÃO, J. (2009). The role of credit officers in the performance of microcredit loans: evidence from VivaCred in Brazil. Mimeo
- AKERLOF, G. A. (1970), "The Market for Lemons: Quality Uncertainty and the Market Mechanism" *Quarterly Journal of Economics* 84 (3): 488–500
- ARROW, K. (1963). "Uncertainty and the Welfare Economics of Medical Care". *American Economic Review* 53 (5): 941–973.
- DIECKMANN, R. (2007). Microfinance: An emerging investment opportunity. *Deutsche Bank Research*.
- GHATAK, M. & GUINNANE, T.W., "The Economics of lending with joint liability: theory and practice." *Journal of Development Economics*, Vol. 60 (1999), pp. 195-228.
- IVATURY, G. & ABRAMS, J., O Mercado para Investimentos Estrangeiros em Microfinanças: Oportunidades e Desafios., *Focus Nota, Consultive Group for Assistance of the Poor*, No. 30, (Ago., 2005)
- LABIE, M., MÉON, P. G. e SZAFARZ, A. (2009). "Discrimination in Microfinance: The Role of Credit Officers". Mimeo
- MISHKIN, F.S. (2009). "The Causes and Propagation of Financial Instability: Lessons for Policymakers" p. 55 – 96. Mimeo
- MORДУCH, J., "The Microfinance Promise." *Journal of Economic Literature*, Vol. 37 (Dec.,1999), pp. 1596-1614.
- RAY, D. (1998). *Development Economics*. Princeton University Press.

STIGLITZ, J. E., & WEISS, A. “*Credit Rationing in Markets with Imperfect Information*”. *The American Economic Review* , Vol. 71 No.3. (Jun.,1981)., pp. 393-410.

The Microbanking Bulletin. *Microfinance Information Exchange*. No. 18 (Spring, 2009)

YUNUS, M.. *O Banheiro dos Pobres*. São Paulo, Editora Ática. (2006)