



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**DEPARTAMENTO DE ECONOMIA APLICADA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**JÁDER GARCIA JUCÁ**

**JOGOS EMPÍRICOS DE ENTRADA: UMA APLICAÇÃO AO SETOR BANCÁRIO  
BRASILEIRO**

**FORTALEZA**

**2017**

JÁDER GARCIA JUCÁ

JOGOS EMPÍRICOS DE ENTRADA: UMA APLICAÇÃO AO SETOR BANCÁRIO  
BRASILEIRO

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Centro de Aperfeiçoamento de Economistas do Nordeste (CAEN) da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Economia.

Orientador: Sérgio Aquino de Souza

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

J84j

Jucá, Jáder Garcia.

JOGOS EMPÍRICOS DE ENTRADA : UMA APLICAÇÃO AO SETOR BANCÁRIO BRASILEIRO /  
Jáder Garcia Jucá. – 2017.  
25 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Economia, Administração,  
Atuária e Contabilidade, Programa de Pós-Graduação em Economia, Fortaleza, 2017.

Orientação: Prof. Dr. Sérgio Aquino de Souza.

1. Economia. 2. Microeconomia. 3. Organização Industrial. 4. Teoria dos Jogos. I. Título.

CDD 330

---

JÁDER GARCIA JUCÁ

JOGOS EMPÍRICOS DE ENTRADA: UMA APLICAÇÃO AO SETOR BANCÁRIO  
BRASILEIRO

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Centro de Aperfeiçoamento de Economistas do Nordeste (CAEN) da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Economia.

Trabalho aprovado em: \_\_\_\_/\_\_\_\_/\_\_\_\_.

---

Prof. Dr. Sérgio Aquino de Souza (UFC)  
Orientador

---

Prof. Dr. Ricardo Antônio de Castro Pereira (UFC)

---

Prof. Dr. Francisco Gildemir Ferreira da Silva (UFC)

FORTALEZA

2017

# Resumo

O setor bancário brasileiro é bastante concentrado, temos que 68% do total dos ativos são detidos por cinco bancos. Outra característica desse setor é a participação dos bancos públicos, ele detêm 42% dos ativos. Este trabalho busca investigar a competição do setor bancário brasileiro ao nível municipal através da metodologia de jogos empíricos. A entrada no setor bancário é modelada como um jogo simultâneo de entrada com informação incompleta. O modelo transforma as ações de competição entre bancos em probabilidades de entrada, mostrando a possibilidade de um banco entrar em um determinado mercado e não a sua entrada de fato. As decisões dos bancos giram em torno do *trade-off* existente entre maior competição e o desejo por certas características do mercado. O modelo é desenvolvido para os seis maiores bancos em ativos e inclui controles regionais para maior robustez. Os resultados mostram que as regiões Norte e Nordeste são as menos atrativas para a atividade bancária, enquanto que a região Sul e Sudeste são as mais atrativas. Outro resultado obtido mostra que os outros bancos se beneficiam da presença da CAIXA, o que não ocorre para o Banco do Brasil e os bancos privados.

**Palavras-chave:** Organização industrial; jogos de informação incompleta; instituições financeiras.

# Abstract

The Brazilian banking sector is very concentrated, we have that 68% of the total assets are held by five banks. Another feature of this sector is the participation of public banks, they held 42% of assets. This paper seeks to investigate the competition of the Brazilian banking sector through the methodology of empirical games. Entry into the banking sector is modeled as a simultaneous entry game with incomplete information. The model transforms actions of competition between banks into entry probabilities, showing the possibility of a bank entering in a given market and not the entry in fact. The decisions of the banks revolve around the existing trade-off between greater competition and the desire for certain market characteristics. The model is developed for the six largest banks in assets and includes controls for greater robustness. The results show that the North and Northeast regions are the less attractive for banking activity, while the South and Southeast are the most attractive. Another result obtained shows that the other banks benefit from the presence of CAIXA, which is not the case for Banco do Brasil and the private banks.

**Keywords:** Industrial organization; games of incomplete information; financial institutions.

# Lista de tabelas

Tabela 1 – Estatística Descritiva das Variáveis Estado . . . . .	12
Tabela 2 – Média das principais variáveis estado de acordo com o número de bancos .	13
Tabela 3 – Presença dos bancos e monopólios . . . . .	14
Tabela 4 – Estimativa dos parâmetros das variáveis estado . . . . .	21
Tabela 5 – Estimativa dos parâmetros de efeito estratégia . . . . .	22

## **Lista de abreviaturas e siglas**

BB	Banco do Brasil
BNDES	Banco Nacional de Desenvolvimento Econômico e Social
MPEC	<i>Mathematical Programs with Equilibrium Constraints</i>
CADE	Conselho Administrativo de Defesa Econômica



# Sumário

	<b>Introdução</b> . . . . .	<b>8</b>
	<b>1 DADOS E ESTATÍSTICA DESCRITIVA</b> . . . . .	<b>11</b>
	<b>2 METODOLOGIA</b> . . . . .	<b>15</b>
<b>2.1</b>	<b>O Modelo Teórico de Entrada</b> . . . . .	<b>15</b>
<b>2.2</b>	<b>O Modelo Empírico de Entrada</b> . . . . .	<b>17</b>
<b>2.2.1</b>	Estágio 1: . . . . .	<b>17</b>
<b>2.2.2</b>	Estágio 2: . . . . .	<b>18</b>
	<b>3 RESULTADOS</b> . . . . .	<b>21</b>
	<b>4 CONCLUSÕES</b> . . . . .	<b>24</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>25</b>

## Introdução

Como é destacado por Bajari et al. (2012) a teoria dos jogos é uma ferramenta bastante aplicada em todos os campos da teoria econômica, especialmente na área de organização industrial. Particularmente nesse campo, a teoria dos jogos mudou a análise das interações de mercado e também serve como uma base importante para a recomendação de políticas públicas. Com isso as análises empíricas de jogos tem tido importância nas pesquisas recentes de econometria e organização industrial.

Nas primeiras aplicações de jogos discretos, os jogos são modelados como problemas de escolha discreta. Porém, as utilidades dos agentes dependem das ações dos outros jogadores. As decisões de entrada das firmas em um mercado deve levar em conta vários aspectos como o tamanho da demanda, os custos, fatores competitivos entre outros aspectos. O que resulta em decisões bastante complexas. Porém, geralmente os pesquisadores só tem acesso, ou só observam, qual decisão a firma tomou. Ou seja, geralmente observamos se a firma está ou não presente em um determinado mercado. O que torna a ferramenta de modelar os jogos como escolhas discretas bastante interessante. Um dos trabalhos pioneiros nesse campo é o de Bresnahan e Reiss (1991).

Como uma extensão do modelo de Bresnahan e Reiss (1991) e aplicação ao setor bancário brasileiro, temos o trabalho de Coelho et al. (2007). Que realiza um estudo sobre os impactos competitivos dos bancos públicos em mercados brasileiros concentrados para o ano de 2000. Diferentemente do modelo de Bresnahan e Reiss (1991) a identidade das firmas entrantes tem importância no modelo, pois o interesse é justamente verificar o impacto da entrada de um banco público em contraste com a entrada de um banco privado na conduta do mercado. Esse modelo tem conclusões que corroboram com as hipóteses feitas pelos autores de que os bancos privados parecem ser mais pró-competitivos do que os bancos públicos.

Na reavaliação do estudo de Coelho et al. (2007) feito por Silva Júnior (2014) para o ano de 2010, os resultados foram semelhantes a respeito de os bancos públicos não afetarem a conduta e a competição com os bancos privados. Porém, é destacado que a exigência do entrante do mercado local tornou-se menor que é, provavelmente, um reflexo de uma diminuição dos custos operativos.

Na literatura mais recente de jogos empíricos tem se destacado o método de dois estágios para estimar jogos de entrada com informação incompleta, como o exposto por Bajari et al. (2012). No primeiro estágio, o economista estima a forma reduzida do modelo. Isso consiste

em usar técnicas de estimação padrão para estimar a probabilidade que uma, dentre um número finito de possíveis escolhas, é observada condicional as covariantes relacionadas. No segundo estágio, o economista estima um *random utility model* de um único agente, incluindo as crenças de equilíbrio do comportamento das outras firmas que é estimado no primeiro estágio. Geralmente é mais fácil estimar jogos discretos em dois estágios.

Além da estimação por dois estágios também há outro método mais recente de estimação desenvolvido por Su e Judd (2012) e explorado por Vitorino (2012). Esse método está incluso nas classes de problemas chamados de *mathematical programs with equilibrium constraints* (MPEC). Diferentemente do método de dois estágios, a estimação via MPEC procede em um só estágio. Basicamente, ele consiste em maximizar uma função de verossimilhança sujeita às restrições de que as variáveis econômicas endógenas sejam consistentes com o equilíbrio definido pelos parâmetros estruturais do modelo.

No dia 2 de agosto de 2015 o Bradesco fez o anúncio formal da compra de todo o aglomerado do HSBC no Brasil por US\$ 5,2 bilhões. O anúncio foi feito em Londres depois de dois meses de negociação envolvendo dois concorrentes, o Santander e o Itaú Unibanco. No dia 5 de janeiro de 2016 o Bradesco anunciou através de um comunicado ao mercado do diretor vice-presidente Alexandre da Silva Glüher que o Banco Central do Brasil aprovou a aquisição de 100% do capital social do HSBC Brasil. Já no dia 8 de junho de 2016 o Conselho Administrativo de Defesa Econômica (CADE) aprovou a aquisição do HSBC condicionada à celebração de um Acordo em Controle de Concentrações, como o objetivo de sanar as preocupações concorrenciais eminentes da aquisição.

Olhando para o cenário do setor bancário ao nível nacional em 2010, podemos concluir que esse mercado pode ser considerado um mercado com alta concentração. De acordo com o Sisbacen, os cinco maiores bancos em ativos que representam 0,28% das instituições financeiras, detinham 68% do total dos ativos. Enquanto que os três maiores, equivalente a 0,17%, detinham 47%. Com respeito aos depósitos totais os cinco maiores bancos detinham 75% e os três maiores 54%. Outro aspecto importante desse mercado é a forte presença dos bancos públicos, eles detinham 42% dos ativos e 45,2% dos depósitos totais. Além disso, o Banco do Brasil (BB), a CAIXA Econômica Federal e o Banco Nacional de Desenvolvimento Econômico e Social (BNDES), que são os três maiores bancos públicos, juntos detinham 42% dos ativos. Um dos pressupostos do modelo abordado neste trabalho é de que a análise é feita em mercados com níveis de concentração elevados, os dados acima corroboram com esse pressuposto.

A figura mostra a elevada concentração do setor financeiro do Brasil ao longo dos anos de 2009 a 2014. O C3 representa os três maiores bancos em ativos e o C5 os cinco maiores, os valores são calculados em relação ao ativo total de todo o sistema financeiro. Podemos verificar que a tendência foi sempre de uma elevação do C3 e C5 nesse período, sendo que o primeiro aumentou em 12% e o segundo em 23%. O C5 em 1995 era composto, em ordem decrescente, pela CAIXA, BB, BANESPA, BRADESCO, BNDES. Retirando-se o BANESPA

que foi comprado pelo Santander, todos os outros quatro bancos acima ainda compõem o C5 em 2014. A partir de 2000 temos o Itaú compondo o C5 no lugar do BANESPA. Em 1995 a CAIXA era o banco com o maior ativo, seguida pelo BB. O que se inverteu em 2000 e continua até em 2014, com o BB sendo o banco de maior ativo.

Figura 1 – Participação do C3 e C5 no Brasil.



Fonte: Banco Central do Brasil.

Além destas considerações iniciais, este trabalho possui quatro seções. A primeira seção apresenta a descrição dos dados utilizados neste trabalho e suas respectivas fontes, assim como as estatísticas descritivas à respeito das variáveis selecionadas. Na segunda seção são descritos o modelo teórico e o empírico. Os resultados da estimação do modelo são relatados e interpretados na terceira seção. Por fim, a última seção destina-se às conclusões.

## 1 DADOS E ESTATÍSTICA DESCRITIVA

Neste trabalho são utilizadas duas bases de dados. A primeira é o Atlas do Desenvolvimento Humano no Brasil que extrai os dados do Censo Demográfico de 2010 do Instituto Brasileiro de Geografia e Estatística. Essa base é utilizada para coletar as informações demográficas dos municípios brasileiros. A segunda base, é a estatística bancária por município do Banco Central do Brasil, chamada de ESTBAN que contém informações sobre o setor bancário ao nível dos municípios. Essa base contém informações sobre o número de agências que cada banco brasileiro tem em cada município brasileiro. Os dados utilizados são referentes ao mês de dezembro de 2010, mês no qual está disponível o censo de 2010.

Para uma melhor adequação ao modelo aqui estudados, é feito um corte nos dados, que consiste em excluir todos os municípios que fazem parte de alguma região metropolitana. Os motivos desse corte são os mesmos apontados por Coelho et al. (2007). Por exemplo, um cliente de um banco que trabalha na cidade principal, ou na capital, pode morar em outra cidade que é parte da mesma região metropolitana. Além disso, quanto mais concentrado for o mercado mais relevantes serão os efeitos competitivos. Assim, cidades pequenas tendem a ter mercados relativamente mais concentrados do que cidades grandes. Portanto, outro corte é feito com base na magnitude da população total em domicílios particulares permanentes (POPT), exceto aqueles com renda nula. Retirando-se todos os municípios que possuem uma população maior que 100,000 e uma população menor do que 10,000, ou seja, a base compreende nos municípios com população de 10,000 a 100,000 habitantes. Os motivos desse corte devem-se ao fato de cidades muito populosas, assim como cidades com população muito pequena, tendem a não se enquadrar no modelo de competição em mercados concentrados. Por fim, temos uma amostra com municípios que provavelmente possuem mercados mais concentrados, sendo justamente o que procuramos modelar neste trabalho. O número total de observações, ou municípios, é reduzido de 5565 para 2521 após os cortes.

Foram selecionados os seis maiores bancos em ativos para a análise deste trabalho, com exclusão do BNDES. Os bancos são: o Banco do Brasil, o Bradesco, a Caixa Econômica Federal, o HSBC, o Itaú Unibanco S.A. e o Santander. Além disso, bancos públicos estaduais também foram incluídos. Os dados e as estatísticas apresentados ao longo deste trabalho serão retirados a partir dessa base de dados que incluem esses bancos e os municípios remanescentes dos cortes descritos acima.

Das características demográficas dos municípios disponíveis foram selecionadas quatro para este trabalho. Elas são denominadas de variáveis estado e são: renda domiciliar per capita (RDPC), população total em domicílios particulares permanentes (POPT), exceto aqueles com renda nula e o índice de Gini. Além das variáveis estados contínuas descritas anteriormente, também são incluídas variáveis categóricas (*dummies*) que incorporam os controles regionais ao modelo. Há duas classes de controles regionais. A primeira é a incorporação de controles por regiões brasileiras que são cinco : Norte, Nordeste, Centro-Oeste, Sul e Sudeste. Para cada região, com exclusão da região Norte, é criada uma variável *dummy* que assume o valor 1 se o município pertence aquela dada região e 0 caso contrário. A região Norte é a variável omitida. Essas variáveis de controle regional tem como principal intuito de captar as diferenças entre os custos de operação entre cada região.

A segunda classe de controles regionais incorpora a presença dos bancos públicos estaduais nos municípios, sendo eles os bancos: Banco da Amazônia (Basa), Banco do Estado de Sergipe (Banese), Banco do Estado do Espírito Santo (Banestes), Banco do Estado do Pará (Banpará) e o Banco do Estado do Rio Grande do Sul (Banrisul). Esse outro controle regional tem o intuito de captar o efeito da presença dos bancos públicos estaduais nas decisões de entrada dos bancos selecionados.

As variáveis estados têm o intuito de explicar a presença, ou a ausência, dos bancos em determinado município. Elas tem o intuito de caracterizar os mercados municipais e impactam diretamente na decisão de entrada dos bancos, essa relação fica mais clara com a descrição do modelo teórico na seção 2.

A estatística descritiva das variáveis estado está contida na tabela 1. A maior parte dos municípios da amostra pertencem à região Nordeste (1070), seguida pela Sudeste (653), Sul (391), Norte (236) e por último a Centro-Oeste (171).

Tabela 1 – Estatística Descritiva das Variáveis Estado

Variável	Média	Desvio Padrão	Máximo	Mediana	Mínimo
RDPC	463.43	234.03	1446.56	398.37	122.21
POPT	25419.84	17632.88	99970.00	19142.00	10001.00
Gini	0.51	0.06	0.80	0.51	0.33
Basa	0.03	0.16	1.00	0.00	0.00
Banrisul	0.05	0.22	1.00	0.00	0.00
Banestes	0.02	0.15	1.00	0.00	0.00
Banpará	0.01	0.09	1.00	0.00	0.00
Banese	0.01	0.12	1.00	0.00	0.00
Norte	0.09	0.29	1.00	0.00	0.00
Nordeste	0.42	0.49	1.00	0.00	0.00
Centro-Oeste	0.07	0.25	1.00	0.00	0.00
Sudeste	0.26	0.44	1.00	0.00	0.00
Sul	0.16	0.36	1.00	0.00	0.00
Observações	2521				

Fonte: Atlas Brasil e Banco Central do Brasil.

A tabela 2 contém as médias da renda per capita, população e renda total de acordo com

o número de diferentes bancos presentes em cada município. Os municípios são divididos em grupos e calcula-se as médias de cada variável para cada grupo. De todos os municípios apenas 130 (5%) possuem agências de seis ou mais bancos, enquanto que 461 (18%) não possuem agências. Podemos ver claramente pela tabela que quanto maior a renda per capita de um município maior será o número de bancos distintos presentes nesse município, o mesmo pode ser constatado para a população. O que corrobora com o fato de a decisão de entrada no setor bancário ser relacionada positivamente com a renda e a população, pois as duas compõem uma medida de tamanho do mercado.

Tabela 2 – Média das principais variáveis estado de acordo com o número de bancos

Nº bancos	Nº observações	Percentual	Renda per capita	População	Renda Total
0	461	18 %	248	14894	3722233
1	751	30 %	349	19039	6120490
2	458	18 %	484	22577	9674034
3	351	14 %	583	28420	14636813
4	235	9 %	665	35004	21453227
5	135	5 %	742	46308	33406917
6	109	4 %	852	15252	52492594
7	21	1 %	554	82696	37338102
Total	2521	100 %	463	25420	12755036

Fonte: Atlas Brasil e Banco Central do Brasil.

Nota: Todas as variáveis estado estão em termos de suas médias. Número de bancos é a quantidade de bancos distintos presentes no município. Renda total é a média da multiplicação da renda per capita e população de cada município. O número máximo de bancos em um município é de 7.

A tabela 2 também ilustra o fato de o setor bancário ser bastante concentrado. Por exemplo, aproximadamente 48% dos municípios possuem um mercado em monopólio ou duopólio. Entretanto, se compararmos apenas os municípios que possuem bancos com os que possuem monopólio ou duopólio, podemos ver que aproximadamente 58% se encontram nessa configuração de mercado.

Na tabela 3 temos representado em quantos municípios cada banco está presente e em quantos municípios ele detém monopólio. O Banco do Brasil é o que mais está presente nos municípios, contando com a presença em 1892 (75%) municípios. O Banco do Brasil também detém o maior número de monopólios 600, sendo este número cerca de sete vezes maior que o número de monopólios do Bradesco, sendo este número o segundo maior no quesito monopólio. O Bradesco também é o segundo mais presente, contando com a presença em 972 (39%) municípios. Desconsiderando os bancos estaduais, teríamos que o banco de menor presença é o HSBC, com presença em 274 (11%) municípios.

Considerando o cenário após a aquisição do HSBC pelo Bradesco em que este decide em manter o mesmo número de agências apenas transformando aquelas que eram do HSBC em Bradesco, teríamos a seguinte configuração: A presença do novo Bradesco seria em 1014 municípios, apenas 42 a mais do que a quantidade que o Bradesco já possuía. Este aumento é

Tabela 3 – Presença dos bancos e monopólios

Banco	Presença	Percentual*	Monopólio
Banco do Brasil	1892	75	600
CAIXA	745	30	7
Bradesco	972	39	83
HSBC	274	11	1
Itaú	584	23	28
Santander	357	14	4
Banrisul	130	5	7
Banestes	58	2	7
Banpará	22	1	2
Banese	34	1	7
Basa	68	3	5
Total	-	-	751

Fonte: Banco Central do Brasil.

Nota: O percentual é calculado em relação ao número total de municípios da amostra.

pequeno pelo fato de que na maioria dos municípios nos quais o HSBC está presente o Bradesco também está. O número de monopólios se alteraria em uma unidade, pois o HSBC possui apenas um monopólio.



## 2 METODOLOGIA

O modelo de entrada explorado neste trabalho, tanto o teórico quanto o empírico, tem como base os modelos desenvolvidos e aplicados por Bajari et al. (2012) e Vitorino (2012). Em primeiro lugar é explicado como funciona o modelo teórico de entrada em seguida é explicado o modelo empírico e sua aplicação no setor bancário brasileiro.

### 2.1 O Modelo Teórico de Entrada

A entrada no setor bancário é modelada como um jogo simultâneo de entrada com informação incompleta. O modelo transforma as ações de competição entre bancos em probabilidades de entrada, mostrando a possibilidade de um banco entrar em um determinado mercado e não a sua entrada de fato. As decisões dos bancos giram em torno do *trade-off* existente entre maior competição e o desejo por certas características do mercado.

No modelo há um número finito de jogadores,  $i = 1, \dots, n$  e cada jogador escolhe simultaneamente uma ação  $a_i \in \{0, 1\}$ . Onde,  $a_i = 1$  se o banco decide entrar no mercado e  $a_i = 0$  caso ele decida não entrar no mercado. Defina  $A = \{0, 1\}^n$  como o vetor de possíveis ações de todos os jogadores e defina  $a = (a_1, \dots, a_n)$  um elemento genérico de  $A$ . Por fim, defina  $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$  como sendo o vetor de estratégias para todos os jogadores com exclusão do jogador  $i$ .

O vetor de variáveis estado do jogador  $i$  no mercado  $m$  é definido por  $s_{im} \in S_{im}$ . Esse vetor inclui variáveis específicas do mercado, sendo elas variáveis contínuas ou discretas. Faça  $S_m = \prod_{i=1}^n S_{im}$  e defina  $s_m = (s_{1m}, \dots, s_{nm}) \in S_m$  como sendo o vetor de variáveis estado de todos os jogadores no mercado  $m$ . É assumido que  $s_m$  é de conhecimento comum para todos os jogadores e para o econometrista.

A variável aleatória  $\varepsilon_i(a_i)$  é um termo de idiosincrasia representando a informação privada da banco  $i$ . A densidade de  $\varepsilon_i = [\varepsilon_i(a_i = 0), \varepsilon_i(a_i = 1)]$  é denotada por  $f(\varepsilon_i)$  e é de conhecimento comum para todos os bancos.

Os lucros *ex post* do banco  $i$  são dados por

$$\tilde{\Pi}_i(a_i, a_{-i}, s_m, \varepsilon_i; \theta) = \Pi_i(a_i, a_{-i}, s_m; \theta) + \varepsilon_i(a_i) \quad (2.1)$$

onde,  $\Pi_i(a_i, a_{-i}, s_m; \theta)$  é uma função conhecida e determinística dos estados  $s_m$ , ações  $a_i$  do

jogador  $i$  e das ações dos outros jogadores  $a_{-i}$ , e  $\theta$  é um vetor de parâmetros.

O *payoff* desse modelo é similar a um *random utility model* como o logit, porém incorpora as ações dos outros jogadores.

Dois pressupostos devem ser feitos para que o modelo seja identificado, esses pressupostos são feitos com base no teorema desenvolvido por Bajari et al. (2012).

Pressuposto 1: Os termos de erro  $\varepsilon_i$  são distribuídos i.i.d valor extremo do Tipo I ao longo das ações  $a_i$  e dos jogadores  $i$ .

Pressuposto 2:  $\Pi_i(a_i = 0, a_{-i}, s_m; \theta) = 0$ .

Como o banco  $i$  não observa os  $\varepsilon_i$  dos outros bancos, ele deve então construir crenças sobre as possíveis ações dos outros bancos, se eles possivelmente entram ou não no mercado, usando toda informação que ele tem disponível. Dado essas crenças, o banco deve então escolher a ação  $a_i \in \{0, 1\}$  que maximiza sua utilidade esperada.

Nesse modelo, a decisão do jogador  $i$  é uma função  $a_i = \gamma(s_m, \varepsilon_i)$ . Defina  $\sigma_i(a_i|s_m)$  como:

$$\sigma_i(a_i = 1|s_m) = \int 1\{\gamma(s_m, \varepsilon_i) = 1\}f(\varepsilon_i)d\varepsilon_i \quad (2.2)$$

onde  $1\{\gamma(s_m, \varepsilon_i) = 1\}$  é uma função indicadora de que a ação do jogador  $i$  é de entrar no mercado. Logo,  $\sigma_i(a_i = 1|s_m)$  é a probabilidade que o jogador  $i$  decide entrar no mercado  $m$  dado  $s_m$  que é de conhecimento comum.

O valor esperado do lucro obtido pelo banco  $i$  associado com a decisão de entrar no mercado é igual a:

$$E[\tilde{\Pi}_i(a_i = 1, a_{-i}, s_m, \varepsilon_i(1); \theta)] = \sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, s_m; \theta)\sigma_{-i}(a_{-i}|s_m) + \varepsilon_i(1) \quad (2.3)$$

onde  $\sigma_{-i}(a_{-i}|s_m) = \prod_{j \neq i} \sigma_j(a_j|s_m)$  representa as crenças que o jogador  $i$  tem sobre as probabilidades de entrada ou não dos outros jogadores. Para cada  $j \neq i$  a probabilidade de o jogador  $j$  entrar no mercado, na perspectiva do jogador  $i$ , é dada por:

$$\sigma_j(a_j = 1|s_m, \theta) = Prob\{E[\tilde{\Pi}_j(a_j = 1, a_{-j}, s_m, \varepsilon_j(1); \theta)] \geq \varepsilon_j(0)\} \quad (2.4)$$

onde a parte determinística do lucro foi normalizada para zero, de acordo com o pressuposto 2.

Como em um equilíbrio de Nash Bayesiano as crenças dos jogadores são iguais as suas probabilidades de escolha, então o equilíbrio de Nash Bayesiano do jogo de entrada estático será uma coleção de crenças  $\sigma_i^*(a_i = 1|s_m, \theta)$  para cada jogador  $i = 1, \dots, n$ . A distribuição de  $\varepsilon$ , de acordo com o pressuposto 1, implica que a probabilidade de o jogador  $i$  escolher entrar no mercado  $m$  é:

$$\sigma_i^*(a_i = 1|s_m, \theta) = \frac{\exp\{\sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, s_m; \theta)\sigma_{-i}(a_{-i}|s_m)\}}{1 + \exp\{\sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, s_m; \theta)\sigma_{-i}(a_{-i}|s_m)\}} \quad (2.5)$$

para  $i = 1, \dots, n$  e  $\sigma_{-i}^*(a_{-i}|s_m) = \prod_{j \neq i} \sigma_j(a_j|s_m)$ .

A probabilidade de equilíbrio  $\sigma^*$  que soluciona o sistema de equações na 2.5 depende das características do mercado  $s_m$  e dos parâmetros estruturais  $\theta$ .

A parte determinística do lucro é especificada como sendo uma função linear das variáveis estado e das ações dos outros jogadores, da forma:

$$\Pi_i(a_i = 1, a_{-i}, s_m, \theta) = \begin{cases} \beta' s_m + \delta \sum_{j \neq i} 1\{a_j = 1\} & \text{se } a_i = 1 \\ 0 & \text{se } a_i = 0 \end{cases} \quad (2.6)$$

onde  $\beta$  mensura o impacto das variáveis estados no lucro de  $i$  e  $\delta$  mensura o impacto da presença de outros bancos no lucro de  $i$ . Temos então que  $\theta = (\beta, \delta)$  é o vetor de parâmetros estruturais. Os parâmetros de efeito estratégia  $\delta$  são os mais importantes desse modelo de entrada. Diferentemente do modelo explorado por Vitorino (2012) em que a parte determinística do lucro de cada jogador é função de um vetor  $\beta_i$  específico do jogador  $i$ , o lucro na equação 2.10 é função de um vetor de parâmetros  $\beta$  que é o mesmo para cada jogador e não um vetor para cada um deles.

De acordo com a especificação do lucro em 2.6 e com os pressupostos 1 e 2, o sistema de equações em 2.5 pode ser escrito como:

$$\sigma_i^*(a_i = 1|s_m, \theta) = \frac{\exp\{\beta' s_m + \delta \sum_{j \neq i} \sigma_j^*(a_j = 1|s_m, \theta)\}}{1 + \exp\{\beta' s_m + \delta \sum_{j \neq i} \sigma_j^*(a_j = 1|s_m, \theta)\}} \quad (2.7)$$

para  $i = 1, \dots, n$  e  $j = 1, \dots, n$ .

## 2.2 O Modelo Empírico de Entrada

O objetivo é estimar os parâmetros estruturais do modelo por meio do equilíbrio descrito na equação 2.7. O método de estimação utilizado é similar ao proposto por Bajari et al. (2012) com pequenas diferenças. Devemos então estimar um modelo semi-paramétrico em dois estágios.

Basicamente, o primeiro estágio consiste em fazer uma estimação não-paramétrica, ou semi-paramétrica,  $\hat{\sigma}_i^*(a_i = 1|s_m)$  das probabilidades de escolha  $\sigma_i^*(a_i = 1|s_m, \theta)$  e em um segundo estágio uma estimação paramétrica para recuperar os parâmetros estruturais.

### 2.2.1 Estágio 1:

Suponha que observamos  $m = 1, \dots, M$  repetições do jogo, ou seja, em cada município  $m$  observamos uma repetição do jogo de entrada. Logo, temos  $M$  observações das ações e variáveis estado de cada jogador  $(a_{i,m}, s_m)$ . Assim, podemos estimar as probabilidades de escolha de cada jogador  $i$  por meio das características do mercado  $s$ . Como destaca Bajari et al. (2012) a

estimativa não-paramétrica do estágio 1 pode ser feita por vários métodos, como *kernel smoothing* ou *local polynomial regressions*.

Neste trabalho é utilizado o *generalized additive model* (GAM). O modelo GAM de escolha discreta é similar aos modelos usuais de escolha discreta como o logit e o probit, porém ao invés de utilizar um preditor linear utiliza-se um preditor não linear para as variáveis contínuas do modelo. Esse preditor consiste em uma soma de funções de alisamento (*smooth functions*) de cada variável explanatória contínua, ou seja, cada variável explanatória contínua se relaciona com a variável dependente por meio da função alisamento estimada para essa variável. Já as variáveis discretas, como as *dummies* entram linearmente no modelo. Assim, temos o seguinte modelo:

$$E[a_i|s_{im}] = f_0 + \sum_k^K f_k(s^k) + \sum_y^Y s^y \quad (2.8)$$

onde  $s^k$  é a variável estado contínua que ocupa a entrada de número  $k$  do vetor  $s_i$  e  $s^y$  é a variável estado discreta que ocupa a entrada de número  $y$  do vetor  $s_i$ . A equação 2.8, de acordo com as variáveis estado utilizadas neste trabalho, pode ser escrita da seguinte forma:

$$\begin{aligned} \sigma_i^*(a_i = 1|s_{im}) = & f_0 + f_1(POPT) + f_2(RDPC) + f_3(Gini) + Nordeste + Sul + Sudeste \\ & + Centro.Oeste + Basa + Banrisul + Banestes + Banpará \\ & + Banese \end{aligned} \quad (2.9)$$

Devemos então estimar a equação 2.9 para cada um dos seis bancos de forma a obter os valores de  $\hat{\sigma}_i^*(a_i = 1|s_m)$  que serão utilizados no estágio 2. Devemos notar que as probabilidades de escolha em 2.9 não dependem das ações dos outros jogadores, apenas das características dos mercados.

## 2.2.2 Estágio 2:

No estágio 2 devemos estimar um modelo paramétrico para poder recuperar o valor do parâmetro estrutural  $\theta$ . Como se trata de um modelo de escolha discreta, devemos então maximizar a seguinte função de verossimilhança:

$$L(\theta) = \prod_{m=1}^M \prod_{j=1}^n [\sigma_j^*(a_j = 1|s_m, \theta)]^{1\{a_j=1\}} [1 - \sigma_j^*(a_j = 1|s_m, \theta)]^{1\{a_j=0\}} \quad (2.10)$$

onde para cada jogador  $i$  temos que  $\sigma_i^*(a_i = 1|s_m, \theta)$  é definido da seguinte forma:

$$\sigma_i^*(a_i = 1|s_m, \theta) = \frac{\exp\{\beta' s_m + \delta \sum_{j \neq i} \hat{\sigma}_j^*(a_j = 1|s_m)\}}{1 + \exp\{\beta' s_m + \delta \sum_{j \neq i} \hat{\sigma}_j^*(a_j = 1|s_m)\}} \quad (2.11)$$

Diferentemente da equação 2.7, na equação 2.11 o lado direito utiliza as estimativas  $\hat{\sigma}_i^*(a_i = 1|s_m)$ , calculadas no estágio 1, para calcular os valores de  $\sigma_i^*(a_i = 1|s_m, \theta)$  que agora incorporam os efeitos estratégia. Assim, podemos estimar  $\theta$  por meio da equação 2.10.

Como o estágio 1 é estimado não-parametricamente devemos esperar que ocorra convergências à uma taxa menor que  $\sqrt{M}$ . Porém, como é mostrado em Bajari et al. (2012) apesar de o estágio 1 ser estimado não-parametricamente os parâmetros estruturais serão assintoticamente normais e irão convergir à uma taxa de  $\sqrt{M}$ . Além disso, sobre condições apropriadas a variância assintótica do estágio 2 será independente do método não-paramétrico que foi escolhido na estimação do estágio 1.

Para melhor análise os bancos são divididos em três grupos. O primeiro grupo contém apenas o BB. O segundo grupo contém apenas a CAIXA e o terceiro grupo contém os bancos privados: Bradesco, HSBC, Itaú e Santander. Portanto, bancos pertencentes ao mesmo grupo vão ter o mesmo valor do parâmetro de efeito estratégia  $\delta$  no *payoff* de outro banco. Por exemplo, os bancos privados terão o mesmo  $\delta$  no *payoff* do BB. Entretanto, como só há o grupo dos bancos privados com mais de um membro, o modelo não sofre muitas alterações. Antes tínhamos 30 parâmetros de efeito estratégia, 5 para cada banco, que com a divisão de grupos ficamos com 7 parâmetros de efeito estratégia. Essa alteração só modifica o modelo no segundo estágio, diretamente na equação 2.11, já que o primeiro estágio não depende dos parâmetros e efeito estratégia  $\delta$ .

Após a estimação de 2.10 devemos realizar um *bootstrap* para computar os desvios padrões dos parâmetros estimados e realizar testes de significância. O processo do *bootstrap* consiste em gerar 500 amostras sintéticas a partir da amostra original. As amostras sintéticas também devem ter o mesmo tamanho da original, ou seja, 2521 observações. Depois da geração das amostras devemos resolver todo o processo de estimação, estágio 1 e 2, para cada nova amostra gerada pelo *bootstrap*. Assim, podemos calcular os desvios padrões dos parâmetros do modelo e realizar testes de significância estatística.

Como é destaque por Canta e Dubois (2015), o processo de geração de amostras sintéticas deve preservar as características dos municípios. Portanto, devemos gerar as amostras sintéticas através de retiradas com reposição dos municípios de tal modo que os mesmos preservem todas as variáveis estados da amostra original. Por exemplo, um determinado município vai aparecer, em cada amostra sintética, com a mesma renda da amostra original, mesmo que este possa aparecer várias vezes, pois o processo é feito com reposição. O que difere do processo mais comum realizado nos *bootstraps* que consiste em retiradas de uma distribuição conjunta de todas as variáveis.

O processo utilizado para computar os testes de significância é o mesmo descrito por Cameron e Trivedi (2005). Estimamos os parâmetros para cada amostra sintética de *bootstrap* e estimamos o desvio padrão de cada parâmetro a partir da distribuição empírica desses parâmetros.

Assim, com a estimação do parâmetro original e seus desvios padrões podemos computar os teste de significância comparando os valores críticos da distribuição de uma normal padrão.

### 3 RESULTADOS

A tabela 4 reporta as estimativas dos parâmetros das variáveis estado e seus respectivos desvios padrões. Dos 13 parâmetros, 11 são estatisticamente significantes sendo 9 ao nível de 1%, um ao nível de 5% e um a 10%. Os sinais positivos dos parâmetros da renda e da população são os esperados pelo modelo e foram confirmados na estimativa. Esses resultados vão de acordo com os obtidos por Coelho et al. (2007) e Silva Júnior (2014), assim como o sinal negativo do parâmetro do índice de Gini, apesar desse não ser estatisticamente significativo.

Tabela 4 – Estimativa dos parâmetros das variáveis estado

Variável	Coefficiente	(Desvio Padrão)
Constante	-5.37*	0.79
POPT	1.13*	0.11
RDPC	0.77*	0.09
Gini	-0.50	0.99
Basa	-0.65***	0.35
Banestes	-2.26*	0.31
Banpará	-1.22**	0.48
Banrisul	-1.84*	0.30
Banese	-0.65	0.50
Nordeste	0.61*	0.19
Centro-Oeste	1.87*	0.27
Sul	2.06*	0.28
Sudeste	2.50*	0.28
Log-Verossimilhança	-5525.289	
Pseudo- $R^2$ ajustado	0.4165	
Observações	2521	

Fonte: Atlas Brasil e Banco Central do Brasil.

Nota: POPT está dividida por 10000 e RDPC por 100.

\* Estatisticamente significativo ao nível de 1%.

\*\* Estatisticamente significativo ao nível de 5%.

\*\*\* Estatisticamente significativo ao nível de 10%.

Com respeito a localização geográfica dos municípios, temos que a região Norte é a menos atrativa para a entrada dos bancos, seguida pela região Nordeste, a Centro-Oeste, a Sul e por último a Sudeste. Sendo as regiões Centro-Oeste e Sudeste com valores próximos dos seus parâmetros. Esses resultados corroboram com os obtidos por Coelho et al. (2007) e Silva Júnior (2014), indicando quais regiões apresentam custos maiores para o funcionamento dos bancos. Como a região Norte é a região menos desenvolvida e tem a pior economia, esperamos um custo fixo maior para tal região. O menor custo fixo está associado a região Sudeste que é a

economicamente mais forte e dinâmica.

Todos os coeficientes das *dummies*, que indicam a presença dos bancos públicos estaduais, apresentam sinal negativo. Esse resultado era de fato o previsto, pois os bancos públicos estaduais deveriam ter impactos negativos nas decisões dos outros bancos. Desses bancos, os mais impactantes são o Banestes e o Banrisul.

Na tabela 5 temos as estimativas dos parâmetros de efeito estratégia e seus respectivos desvios padrões. Desses 7 parâmetros, apenas 1 não é estatisticamente significativo.

Tabela 5 – Estimativa dos parâmetros de efeito estratégia

	Banco do Brasil	CAIXA	Bancos Privados
Banco do Brasil	-	5.09* (1.92)	-0.46 (0.69)
CAIXA	-3.10* (0.37)	-	-0.76* (0.26)
Bancos Privados	-3.18* (0.34)	0.93** (0.37)	-2.27* (0.27)

Fonte: Atlas Brasil e Banco Central do Brasil.

Nota: Desvios padrões em parênteses.

\* Estatisticamente significativa ao nível de 1%.

\*\* Estatisticamente significativa ao nível de 5%.

\*\*\* Estatisticamente significativa ao nível de 10%.

A tabela 5 pode ser interpretada da seguinte forma: cada linha corresponde a decisão de entrada do banco nela indicado e os valores contidos na linha são os valores dos parâmetros de efeito estratégia correspondentes aos bancos, ou grupo, indicados por cada coluna. Portanto, olhando para a linha do BB, temos que o primeiro valor corresponde ao impacto da CAIXA na decisão de entrada do BB, enquanto que o segundo corresponde ao impacto dos bancos privados. Podemos também analisar a partir de cada coluna, onde cada valor corresponde ao efeito estratégia do banco de cada coluna na decisão de entrada dos bancos em cada linha.

O sinal negativo dos parâmetros de efeito estratégia é o esperado pelo modelo para todos os bancos, porém temos dois parâmetros com sinais positivos e que são estatisticamente significantes. O primeiro desses parâmetros é o efeito da CAIXA no BB que é de 5,09, o segundo é o efeito da CAIXA no grupo dos bancos privados que é de 0.93. Esses resultados indicam que os outros bancos e a CAIXA são estrategicamente complementares, isto é, tanto o BB quanto os bancos privados se beneficiam da presença da CAIXA. Entretanto, o contrário não é verdade, o BB e os bancos privados impactam negativamente na CAIXA. Esse resultado reforça as conclusões de Coelho et al. (2007) e Silva Júnior (2014) à respeito da presença dos bancos públicos terem efeitos neutros na conduta, porém, valendo apenas para a CAIXA.

Como apenas o grupo dos bancos privados tem mais de um banco, a diagonal da tabela 5 só contém um valor. O valor negativo desse parâmetro indica que os bancos privados são estrategicamente substitutos.



Com relação ao BB, verificamos que ele tem o maior efeito estratégia na CAIXA e nos bancos privados e esses valores são negativos e estatisticamente significantes. Indicando que o BB, apesar de ser um banco público, tem mais impacto concorrencial nos bancos privados do que os próprios bancos privados. Diferentemente da CAIXA, os resultados obtidos para o BB sugerem que ele se comporta mais como um banco privado do que um banco público. Talvez isso se dê ao fato de o BB ser um banco público de economia mista tendo, portanto, um comportamento diferente da CAIXA.

Os resultados obtidos para a CAIXA e BB também corroboram com os obtidos por Silva Júnior (2014) para o estudo realizado considerando a CAIXA como o único banco público. Neste trabalho obtemos resultados similares, porém, não consideramos o comportamento do BB e da CAIXA diferente do comportamento dos bancos privados. Além disso, foi possível analisar que apesar de a CAIXA não impactar negativamente nas decisões de entrada dos outros bancos, o contrário acontece.

## 4 CONCLUSÕES

Estimamos um jogo simultâneo de entrada com informação incompleta através do método de dois estágios desenvolvido por Bajari et al. (2012), para o setor bancário brasileiro. No primeiro estágio foi estimado um modelo semi-paramétrico da classe de modelos GAM. No segundo estágio foi estimado um modelo paramétrico através do método de máxima verossimilhança. O modelo incorpora controles regionais de forma a obter diferenciações dos mercados ao nível regional.

Os sinais obtidos para as variáveis estado foram os previstos e a maioria foi estatisticamente significativa. Os resultados mostram que há diferenças regionais, sendo a região Norte a menos atrativa e a região Sudeste a mais atrativa. Além disso, a presença dos bancos públicos estaduais tem impactos negativos nas decisões de entrada dos bancos analisados nesse trabalho.

A respeito dos parâmetros de efeito estratégia, concluímos que a CAIXA tem um efeito positivo na decisão de entrada do BB e do grupo dos bancos privados. O BB é o que tem o maior efeito negativo na CAIXA e nos bancos privados. Já os bancos privados tem efeito negativo em outros bancos privados, na CAIXA e um pequeno efeito no BB. Os resultados sugerem que o BB e a CAIXA tem comportamentos diferentes, sendo o primeiro se assemelhando ao comportamento dos bancos privados enquanto que o segundo não.

## REFERÊNCIAS

- BAJARI, P. et al. Estimating static models of strategic interactions. *Journal of Business & Economic Statistics*, Taylor & Francis, 2012.
- BRESNAHAN, T. F.; REISS, P. C. Empirical models of discrete games. *Journal of Econometrics*, Elsevier, v. 48, n. 1, p. 57–81, 1991.
- CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics: methods and applications*. [S.l.]: Cambridge university press, 2005.
- CANTA, C.; DUBOIS, P. Smoking within the household: Spousal peer effects and children's health implications. *The BE Journal of Economic Analysis & Policy*, v. 15, n. 4, p. 1939–1973, 2015.
- COELHO, C. A. et al. Are public banks pro-competitive? evidence from concentrated local markets in brazil. *Textos para discussão*, v. 551, 2007.
- SILVA JÚNIOR, R. F. d. *Houve alteração na Competitividade de Bancos Públicos e Privados em Mercados Locais Brasileiros após o ano 2000? Uma análise concorrencial para o ano de 2010*. Tese (Doutorado em Economia). p.44 Universidade Federal do Ceará, 2014.
- SU, C.-L.; JUDD, K. L. Constrained optimization approaches to estimation of structural models. *Econometrica*, Wiley Online Library, v. 80, n. 5, p. 2213–2230, 2012.
- VITORINO, M. A. Empirical entry games with complementarities: An application to the shopping center industry. *Journal of Marketing Research*, American Marketing Association, v. 49, n. 2, p. 175–191, 2012.