

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

EDUARDO SOARES CALAZANS

ESTRUTURA DE MERCADO: UMA ANÁLISE DO IMPACTO DA
PRESENÇA ASCENDENTE DE BANCOS 100% DIGITAIS NA
INDÚSTRIA BANCÁRIA BRASILEIRA.

PONTA GROSSA
2022

EDUARDO SOARES CALAZANS

**ESTRUTURA DE MERCADO: UMA ANÁLISE DO IMPACTO DA
PRESENÇA ASCENDENTE DE BANCOS 100% DIGITAIS NA
INDÚSTRIA BANCÁRIA BRASILEIRA.**

Dissertação apresentada como requisito parcial
à obtenção do título de Mestre em Ciências
Econômicas do Programa de Pós-Graduação em
Economia da Universidade Estadual de Ponta
Grossa.

Orientador: Prof. Dr. Armando Vaz Sampaio

**PONTA GROSSA
2022**

C143 Calazans, Eduardo Soares
Estrutura de mercado: uma análise do impacto da presença ascendente de bancos 100% digitais na indústria bancária brasileira. / Eduardo Soares Calazans. Ponta Grossa, 2022.
102 f.

Dissertação (Mestrado em Economia - Área de Concentração: Economia), Universidade Estadual de Ponta Grossa.

Orientador: Prof. Dr. Armando Vaz Sampaio.

1. Indústria bancária. 2. Competição (Panzar-Rosse). 3. Concentração (Herfindahl-Hirschman). 4. Bancos digitais. 5. Análise exploratória de dados espaciais. I. Sampaio, Armando Vaz. II. Universidade Estadual de Ponta Grossa. Economia. III.T.

CDD: 332.66



UNIVERSIDADE ESTADUAL DE PONTA GROSSA
Av. General Carlos Cavalcanti, 4748 - Bairro Uvaranas - CEP 84030-900 - Ponta Grossa - PR - <https://uepg.br>

TERMO

TERMO DE APROVAÇÃO

EDUARDO SOARES CALAZANS

“ESTRUTURA DE MERCADO: UMA ANÁLISE DO IMPACTO DA PRESENÇA ASCENDENTE DE BANCOS 100% DIGITAIS NA INDÚSTRIA BANCÁRIA BRASILEIRA”

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Economia, da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora:

Ponta Grossa, 30 de março de 2022

Armando Vaz Sampaio – Universidade Estadual de Ponta Grossa

Alex Sander Souza do Carmo - Universidade Estadual de Ponta Grossa

Maurício Vaz Lobo Bittencourt - Universidade Federal do Paraná



Documento assinado eletronicamente por **Alex Sander Souza do Carmo, Coordenador(a) do Programa de Pós-Graduação em Economia - Mestrado**, em 31/03/2022, às 11:36, conforme Resolução UEPG CA 114/2018 e art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **Mauricio Vaz Lobo Bittencourt, Usuário Externo**, em 31/03/2022, às 14:20, conforme Resolução UEPG CA 114/2018 e art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **Armando Vaz Sampaio, Professor(a)**, em 02/04/2022, às 08:06, conforme Resolução UEPG CA 114/2018 e art. 1º, III, "b", da Lei 11.419/2006.

A autenticidade do documento pode ser conferida no site <https://sei.uepg.br/autenticidade> informando



o código verificador **0909935** e o código CRC **2F8B04B6**.

22.000023073-9

0909935v4

AGRADECIMENTOS

A Deus, por todas as bênçãos recebidas.

Ao meu orientador, Professor Doutor Armando Vaz Sampaio, pelo direcionamento e crença na minha capacidade de executar as tarefas que este trabalho exigiu.

Ao corpo docente, pelos ensinamentos transferidos, pela dedicação e disponibilidade em tempos tão difíceis.

Aos membros da banca Professor Doutor Armando Vaz Sampaio, Professor Doutor Alex Sander Souza e Professor Doutor Mauricio Vaz Lobo Bittencourt pelas observações, críticas, apontamentos e sugestões feitas ao trabalho.

Aos colegas de mestrado, pela remota, mas excelente parceria.

Em especial, agradeço ao Professor Doutor Alex Sander Souza e a Professora Doutora Luma de Oliveira, por todas as conversas, e-mails respondidos, sugestões, ensinamentos, indicações de livros e artigos. Sem o auxílio de vocês, tenho certeza de que o caminho seria muito mais complicado.

Não poderia deixar de agradecer muito a minha companheira Professora Doutora Camilla Oliveira, por me incentivar, por acreditar, por me apoiar, por todas as horas de dedicação em todos os aspectos, por todo o carinho e por toda a compreensão. Muito obrigado.

Agradeço muito ao meu irmão André, por todas as correções ortográficas feita em quase todos os trabalhos que eu fiz durante este mestrado. Pode escolher a cerveja, eu encho os *growlers*.

E finalmente a minha família, pai, mãe, irmãos sobrinhos, sobrinha e todos que me dão suporte para que eu alcance este objetivo.

Muito obrigado!!!

RESUMO

A presente dissertação tem como objetivo analisar o impacto do crescimento dos bancos digitais sobre a estrutura de mercado da indústria bancária brasileira sob a ótica da concentração, utilizando o Índice de *Herfindahl-Hirschman* normalizado, e da competição, empregando a metodologia da estatística H de Panzar-Rosse. Também é elaborada uma breve discussão sobre a distribuição das agências bancárias no território brasileiro, e uma análise resumida das regulamentações promovidas pelo Banco Central para impulsionar a desconcentração e promover a competição neste setor. As variáveis utilizadas foram dispostas em um painel não balanceado contendo 159 instituições financeiras atuantes no Brasil entre o primeiro trimestre de 2000 ao quarto trimestre de 2019. Os resultados obtidos apontam para um mercado com características de concorrência monopolista e níveis moderados de concentração, além de uma perspectiva de melhoras destes aspectos, dado o crescimento do número de bancos digitais no mercado bancário brasileiro.

Palavras-chaves: Concentração. Competição. Indústria Bancária. Bancos Digitais.

ABSTRACT

The present dissertation aims to analyze the impact of the digital banks growth on the market structure of the Brazilian banking industry from the perspective of concentration, using the normalized *Herfindahl-Hirschman* Index, and competition, using the methodology of H statistic from Panzar-Rosse. It is also elaborated a brief discussion on the distribution of bank agencies in the Brazilian territory, as well as a summarized analysis of the regulations promoted by the Central Bank to promote deconcentration and encourage competition in this sector. The variables used were arranged in an unbalanced panel containing 159 financial institutions operating in Brazil between the first quarter of 2000 and the fourth quarter of 2019. The results obtained point out to a market with characteristics of monopolistic competition and moderate levels of concentration, in addition is show a prospect of improvements on the concentration and competition levels, given to the growth number of digital banks in the market.

Keywords: Concentration. Competition. Banking Industry. Digital Banks.

LISTA DE FIGURAS

Figura 1 – Quantidade de agências bancárias no Brasil entre o primeiro trimestre de 2000 e o quarto trimestre de 2019.	14
Figura 2 – Número de bancos digitais no Brasil entre o primeiro trimestre de 2000 e o quarto trimestre de 2019.	16
Figura 3 – Comportamento temporal da variável Receita da Intermediação Financeira da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	43
Figura 4 – Oscilação anual da variável Receita da Intermediação Financeira da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	44
Figura 5 – Comportamento temporal da variável Depósito Total da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	46
Figura 6 – Oscilação anual da variável Depósito Total da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	46
Figura 7 – Distribuição espacial das agências bancárias e da densidade demográfica dos municípios brasileiros em dezembro de 2019.	54
Figura 8 – Distribuição espacial das operações de crédito e dos depósitos em poupança dos municípios brasileiros em dezembro de 2019.	56
Figura 9 – Distribuição espacial do número de instituições bancárias e, da estimativa populacional brasileiros em dezembro de 2019.	58
Figura 10 – Distribuição espacial do número de instituições bancárias e, da estimativa populacional brasileiros em dezembro de 2019.	60
Figura 11 – $IHHn$ dos Depósitos Totais entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	63
Figura 12 – $IHHn$ dos Saldos das Operações de Crédito entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	64
Figura 13 – $IHHn$ dos Ativos Totais Ajustados entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.	65
Figura 14 – $IHHn$ das variáveis Depósito Total, Saldo das Operações de Crédito e Ativo Total Ajustado do primeiro trimestre de 2015 ao quarto trimestre de 2019.	66
Figura 15 – Gráfico de dispersão das médias trimestrais do Custo Médio e em relação a Receita da Intermediação Financeira.. . . .	67
Figura 16 – Curva da <i>proxy</i> de Custo Médio e Preço do Crédito.	68

Figura 17 – Gráfico de dispersão das médias trimestrais do custo médio e <i>proxy</i> do preço do crédito.	69
Figura 18 – Gráfico de dispersão e histograma da estimativa da estatística H^{RIF} de Panzar-Rosse entre o segundo trimestre de 2000 até o quarto trimestre de 2019.	77
Figura 19 – Gráfico de dispersão e histograma da estimativa da estatística H^{RT} de Panzar-Rosse entre o segundo trimestre de 2000 até o quarto trimestre de 2019.	78

LISTA DE TABELAS

Tabela 1 – Representatividade das amostras, filtrada para a classificação b1 e filtrada para a classificação b1 com tratamento, em comparativo a indústria bancária brasileira nos anos de 2017, 2018 e 2019.	33
Tabela 2 – Estatística descritiva das variáveis	41
Tabela 3 – Estatística QS da variável Receita da Intermediação Financeira da instituição código 10045.	45
Tabela 4 – Estatística QS da variável Depósito Total da instituição código 10045.	47
Tabela 5 – Testes de Raiz Unitária das variáveis RIF e DT e w da instituição código 10045.	47
Tabela 6 – Estatística descritiva das variáveis dessazonalizadas, com transformação logarítmica e os números ausentes tratados.	50
Tabela 7 – Matriz de correlação das variáveis tratadas, dessazonalizadas e loglinearizadas.	51
Tabela 8 – Análise dos 10 municípios brasileiros com a maior densidade populacional em 2019.	53
Tabela 9 – Análise comparativa de dados bancários com a população estimada por estados para o ano de 2019.	61
Tabela 10 – Regressão - Painel, Equação da receita na forma reduzida.	71
Tabela 11 – Variância Robusta das especificações H^{RIF} e H^{RT}	75
Tabela 12 – Variância Robusta	76
Tabela 13 – Intervalo de confiança da distribuição trimestral das estimações das estatísticas H^{RIF} e H^{RT} entre o segundo trimestre de 2000 até o quarto trimestre de 2019	78
Tabela 14 – Variáveis que apresentaram sazonalidade	86
Tabela 15 – Valores estimados do IHH_n , do $\overline{CM}$, do $\overline{PC}$ e da Estatística H	99

SUMÁRIO

1	INTRODUÇÃO	12
2	REFERENCIAL TEÓRICO	18
2.1	DEFINIÇÕES	18
2.2	ORGANIZAÇÃO INDUSTRIAL EMPÍRICA	19
2.2.1	Modelos de competição - resumo	23
2.3	CONCENTRAÇÃO E CONCORRÊNCIA NA INDÚSTRIA BANCÁRIA BRASILEIRA	25
2.4	EVOLUÇÕES NORMATIVAS DO BANCO CENTRAL DO BRASIL	29
3	METODOLOGIA	31
3.1	DADOS	31
3.2	VARIÁVEIS	33
3.3	MODELO EMPÍRICO PROPOSTO DA ESTATÍSTICA H DE PANZAR-ROSSE	37
3.4	ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS	39
3.5	TRATAMENTO DOS DADOS	42
3.5.1	Sazonalidade	42
3.5.2	Raiz unitária	47
3.5.3	Normalização logarítmica	48
3.5.4	Dados faltantes	48
3.5.5	Matriz de correlação	51
4	RESULTADOS	52
4.1	ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS	52
4.1.1	Análise municipal	52
4.1.2	Análise estadual	57
4.2	ÍNDICE DE <i>HERFINDAHL-HIRSCHMAN</i> NORMALIZADO	62
4.2.1	IHH_n com e sem os bancos digitais	65
4.3	CURVAS DE CUSTO MÉDIO E PREÇO DO CRÉDITO	67
4.4	ESTATÍSTICA H DE PANZAR-ROSSE	69
4.4.1	Testes dos modelos	72
4.4.2	Inferência Estatística H de Panzar-Rosse	77
5	CONSIDERAÇÕES FINAIS	80
	REFERÊNCIAS	82
	APÊNDICE A – TABELA DOS TESTES DE SAZONALIDADE	86

APÊNDICE B – DIAGNÓSTICO DO PRÉ-AJUSTE E O MO- DELO ARIMA	89
APÊNDICE C – DIAGNÓSTICOS DOS TESTES <i>DICKEY-FULLER</i> AUMENTADO	90
APÊNDICE D – TRATAMENTO DOS DADOS APÓS O FIL- TRO DE SAZONALIDADE E A TRANSFOR- MAÇÃO LOGARÍTMICA	98
APÊNDICE E – ESTIMATIVAS TRIMESTRAIS DO ÍNDICE DE <i>HERFINDAHL-HIRSCHMANN</i> NORMALIZADO, DO CUSTO MÉDIO, DA <i>PROXY</i> DO PREÇO DO CRÉDITO E DA ESTATÍSTICA <i>H</i>	99
APÊNDICE F – DIAGNÓSTICO DOS TESTES DE DEPENDÊN- CIA TRANSVERSAL, DE <i>BREUSH-PAGAN</i> PARA HETEROCEDASTICIDADE E DE CORRELA- ÇÃO SERIAL EM PAINEL	101

1 INTRODUÇÃO

As constantes mudanças tecnológicas têm impactado intensamente o mundo e, consequentemente, o mercado bancário. Graças aos novos desenvolvimentos em programação e análise de dados, à democratização do acesso à internet com crescente oferta dos serviços domiciliares e empresariais e à ascensão de serviços móveis, a indústria bancária se vê forçada a evoluir digitalmente. Lipton, Shrier e Pentland (2016) pontuam que por anos as novas tecnologias têm sido adicionadas de forma a incrementar as operações existentes, contudo, não no desenvolvimento de novas tecnologias e serviços, mas na sobreposição ou extensão secundária da infraestrutura bancária preestabelecida. Este cenário está mudando drasticamente via demanda de uma clientela insatisfeita com os serviços engessados prestados pelas firmas tradicionais, via regulamentações e desregulamentações que permitem e incentivam a entrada de novas firmas na indústria, via o aperfeiçoamento dos protocolos de segurança de dados e via as inovações tecnológicas.

Neste contexto, os bancos 100% digitais surgem na indústria ofertando contas sem tarifas, maior rentabilidade e serviços mais atrativos que as firmas clássicas. É importante lembrar que nem todo banco denominado 100% digital – doravante chamado apenas bancos digitais – é uma empresa que acabou de nascer. No cenário brasileiro, a maioria destas firmas são instituições financeiras que, ou migraram para o mercado de bancos comerciais, ou alteraram toda sua estratégia empresarial para o modelo de negócio baseado no atendimento remoto. Verifica-se que esta tendência é embalada por uma série histórica de regulamentações que tanto balizam e autorizam a atuação das *fintechs*¹, como estimulam a digitalização bancária.

É preciso lembrar que, embora as instituições financeiras da atualidade ofereçam um leque diferenciado de serviços e produtos, o objetivo principal da atividade bancária permanece a intermediação financeira. “No nível microeconômico, os bancos podem ser entendidos como máquinas produtoras de dividendos que buscam depósitos e emitem empréstimos. No nível macroeconômico, são criadores de dinheiro de crédito.” (LIPTON; SHRIER; PENTLAND, 2016, p. 3, tradução nossa)².

Quando se fala sobre a concorrência na indústria bancária, é importante informar que há um grande debate sobre *trade-off* entre risco e competição. Neto, Araújo e Ponce (2006) argumentam que, em um ambiente de concorrência, na competição pela concessão de créditos, os bancos tendem a flertar com perfis não aptos a honrarem com seus

¹ “As fintechs são reconhecidas pelo Banco Central do Brasil como um grupo de instituições cujos modelos de negócio se baseiam fortemente em tecnologia inovadora e na utilização intensiva de canais eletrônicos de atendimento e de processamento de serviços.” (BANCO CENTRAL DO BRASIL, 2019b, p. 142)

² “At micro level, banks can be thought of as dividend producing machines seeking deposits and issuing loans. At macro level, they are creators of credit money.” (LIPTON; SHRIER; PENTLAND, 2016, p. 3)

compromissos financeiros, expondo-se a um elevado nível de risco e, consequentemente, reduzindo a estabilidade do sistema. Por outro lado, a maior competição pode ser benéfica para a alocação dos recursos na economia, para conduzir o mercado a preços de crédito menores e para elevar o bem-estar social. Belleflame e Peitz (2010) pontuam que o postulado principal para as políticas de incentivo a concorrência é que a concorrência é uma força essencial para proporcionar eficiência econômica.

Apesar da relevância e dicotomia do tema, não pacífico na literatura, observa-se neste trabalho que as autoridades reguladoras no Brasil estão evidenciando fortes sinais de interesse em avançar na formulação de normas e procedimentos que garantam maior segurança jurídica e estimulem a competição no sistema financeiro. Independente do ambiente de competição ou do nível de concentração, Kadir et al. (2014) ressaltam que a performance da indústria bancária precisa ser monitorada. No intuito de prevenir colapsos financeiros, crises sistêmicas e orientar boas práticas, no Brasil, os responsáveis pelo monitoramento e regulamentação do sistema financeiro são o Banco Central (BACEN), o Conselho de Valores Mobiliários e o Ministério da Economia.

No ensejo de monitorar, compreender estruturas de mercado e avaliar o comportamento das firmas, a Organização Industrial Empírica usa de combinação de dados, modelos e métodos econométricos. Este ensaio, que tem como objetivo analisar o impacto do crescimento dos bancos digitais sobre a estrutura de mercado da indústria bancária brasileira, utiliza, para tal, uma técnica da abordagem estrutural – Índice de *Herfindahl-Hirschman* normalizado (IHH_n) – para estimar o nível de concentração e uma técnica da abordagem não estrutural – Estatística H de Panzar e Rosse – para estimar o nível de competição do mercado.

Estudar, analisar e monitorar a estrutura de mercado da indústria bancária se faz necessário uma vez que “os bancos desempenham um papel fundamental no fornecimento de crédito, no sistema de pagamentos, na transmissão da política monetária e na manutenção da estabilidade financeira.” (BIKKER; SHAFFERS; SPIERDIJK, 2012, p. 1025, tradução nossa)³

Em suma, um sistema financeiro eficiente: i) aloca os recursos, poupança e expedientes financeiros, de forma mais eficiente, ii) tende a mitigar assimetrias de informações proporcionando custos de crédito menores, iii) provoca aumento da liquidez econômica, o que resulta em custo de transações menores, para fomentar e viabilizar os investimentos produtivos, iv) oferece proteção, diversificação ou arranjos de combinação do risco enfrentado pelos agentes econômicos, v) facilita e agiliza a troca de bens e serviços em uma economia de mercado, vi) proporciona um melhor monitoramento da qualidade do investimento e, vii) negocia com um maior número de tomadores e emprestadores, diversificando as operações. Logo, um sistema financeiro eficiente é desejável por ser um dos

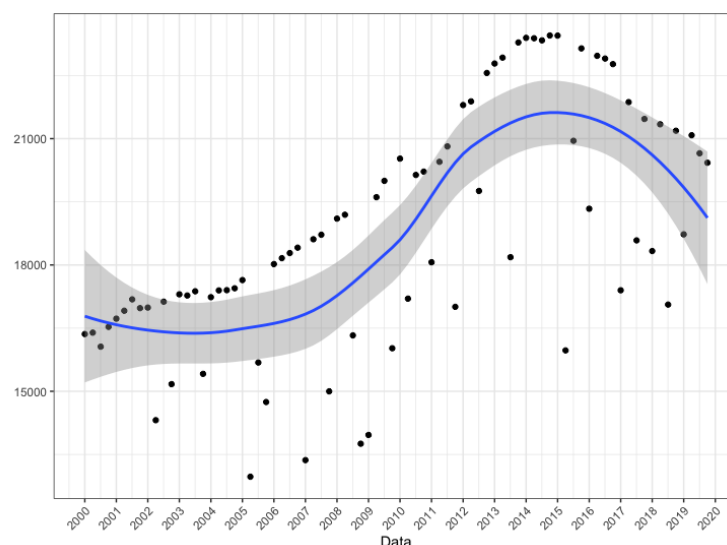
³ “Banks play a pivotal role in the provision of credit, the payment system, the transmission of monetary policy, and maintaining financial stability.” (BIKKER; SHAFFERS; SPIERDIJK, 2012, p. 1025)

fatores determinantes do crescimento econômico.

Neste sentido, diversos ensaios empíricos foram realizados com o objetivo de verificar o ambiente de competitividade do sistema bancário. Busca-se, com este estudo, analisar o impacto do crescimento dos bancos digitais sobre a estrutura de mercado da indústria bancária brasileira sob a ótica da concentração, utilizando o Índice de *Herfindahl-Hirschman* normalizado, e da competição, empregando a metodologia da estatística *H* de Panzar-Rosse, a partir dos dados dos conglomerados financeiros e instituições individuais no intervalo temporal compreendido entre o primeiro trimestre do ano de 2000 até o quarto trimestre do ano de 2019.

Ademais, é um objetivo específico deste ensaio apresentar uma breve discussão sobre a distribuição das agências bancárias no território brasileiro, uma vez que o BANCO CENTRAL DO BRASIL (2019b) aponta a redução de 10,5% no número de agências bancárias no país desde 2014. O BANCO CENTRAL DO BRASIL (2019b) ressalta que o declínio na quantidade de agências bancárias tende a se acentuar com o crescimento e a melhora na oferta de serviços remotos. Com isso, os bancos “tradicionais” tendem a migrar suas operações para a modalidade remota, ofertando cada vez mais produtos isentos de atendimento físico e, conseqüentemente, menos produtos e serviços presenciais. O gráfico da figura 1 deixa nítido o ponto de inflexão da curva de tendência. Observa-se com clareza a reversão da tendência de crescimento da quantidade de agências bancárias em 2014.

Figura 1 – Quantidade de agências bancárias no Brasil entre o primeiro trimestre de 2000 e o quarto trimestre de 2019.



Fonte: Banco Central do Brasil.

Nota: A linha azul é calculada via Mínimos Quadrados Ordinários, o sombreado cinza que envolve a linha azul representa o intervalo de confiança de 95%. A amostra utilizada contém 159 códigos de Conglomerados Financeiros e Instituições Independentes filtrados para o Tipo de Consolidado Bancário b1.

Neste sentido, é importante ressaltar que as regiões menos densas demograficamente tendem a sofrer ainda mais com a redução das agências bancárias, haja vista que estas mesmas regiões são também as que mais carecem da infraestrutura tecnológica para o acesso remoto. Mas, por outro lado, a infraestrutura necessária também está em franca ascensão, graças a demanda de praticamente todos os setores da economia por conexões através da internet.

Apesar dos esforços dos bancos “tradicionais”, Lipton, Shrier e Pentland (2016) pontuam que o crescimento dos bancos digitais⁴ está diretamente ligado à insatisfação dos clientes com os serviços e produtos engessados das firmas “tradicionais”, e não com a quantidade ou presença geográfica das firmas. Tal comportamento abre espaço para que novas firmas apareçam com objetivo de atender a este nicho. Além da insatisfação supracitada, a inclusão de medidas de incentivo a concorrência – implementadas pelos órgãos regulatórios – favorece ainda mais a entrada de novas empresas na indústria.

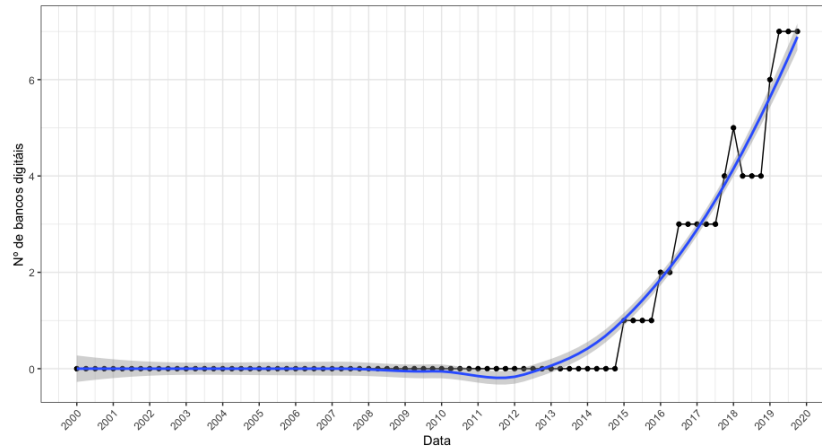
Até dezembro de 2019, dentro do segmento bancário comercial (b1⁵), 8 instituições são classificadas como bancos digitais.⁶ É possível ver no gráfico da figura 2, como é significativo o crescimento desta modalidade de firmas na indústria bancária brasileira.

⁴ A definição de bancos digitais utilizada é a do (BANCO CENTRAL DO BRASIL, 2019b, p 147)

⁵ b1 – Banco Comercial, Banco Múltiplo com carteira comercial ou Caixa Econômica.

⁶ 1) Banco Inter, antigo Banco Intermedium, lança a conta digital em 2015 e o aplicativo para telefonia celular em 2016. É considerado digital a partir do primeiro trimestre de 2015. 2) Original, antes era uma instituição financeira pertencente ao Grupo Matone, lançou a plataforma digital em 2016. É considerado digital a partir do primeiro trimestre de 2016. 3) Neon, é o antigo Banco Pottencial, não há informação de quando virou digital, é considerado a data da alteração do contrato social em que é alterado também no nome social para NEON no terceiro trimestre de 2016. Não há dados do banco na categoria b1 a partir do primeiro trimestre de 2018. 4) BS2, faz parte do grupo Bonsucesso. O Banco digital foi criado no quarto trimestre de 2017. 5) Agibank, foi Banco Agiplan até 12/2017, virou banco digital e alterou o nome no primeiro trimestre de 2018. Não há informações sobre quando começou a ser digital, é considerado a data de alteração do contrato social que deu origem ao Agibank. 6) C6 Bank, único banco desta lista que nasceu digital. Começa as atividades no primeiro trimestre de 2019. 7) Omini, informa que é banco digital desde o primeiro trimestre de 2019. 8) Smartbank, pertence ao grupo Indusval, é considerado digital a partir do segundo trimestre de 2019, data da alteração no contrato social que dá origem ao Smartbank.

Figura 2 – Número de bancos digitais no Brasil entre o primeiro trimestre de 2000 e o quarto trimestre de 2019.



Fonte: Banco Central do Brasil.

Nota: A linha azul é calculada via Mínimos Quadrados Ordinários, o sombreado cinza que envolve a linha azul representa o intervalo de confiança de 95%. A amostra utilizada destaca somente os bancos digitais.

Cada ponto, no gráfico da figura 2, indica a quantidade de bancos digitais para cada trimestre, entre o primeiro trimestre de 2015 e o quarto trimestre de 2019. A linha preta é utilizada para guiar os olhos ao movimento das quantidades. Observa-se que o crescimento é ascendente (de 0 unidades até 7) a partir do primeiro trimestre de 2015 até o quarto trimestre de 2019.

Um dos fatores reconhecidos a partir de reformulações do trabalho de Solow (1956) é o papel do sistema financeiro e a sua relação com o crescimento econômico. É observado que países com um sistema financeiro bem desenvolvido são capazes de promover alto crescimento econômico através da inovação tecnológica, que por sua vez induz ao aumento na demanda por produtos e serviços financeiros. Em resposta a essas demandas, tais mudanças tendem a induzir maior crescimento econômico. Neste sentido, o segundo objetivo específico da presente dissertação é apresentar uma análise resumida das regulamentações promovidas pelo Banco Central para impulsionar a competição, bem como para balizar e viabilizar o ambiente institucional a fim de que o serviço bancário digital se desenvolva.

No ensejo de dar fluência à leitura e organizar de forma objetiva as ideias, este trabalho está dividido em seções e subseções, a começar por esta Introdução. Na segunda seção são apresentadas uma breve revisão da literatura e as regulamentações que balizam a hipótese de que o BACEN vem agindo no sentido de reduzir os níveis de concentração e promover a competição no setor bancário. Na terceira seção serão expostos os dados, as variáveis, a estatística descritiva das variáveis, a metodologia do *IHHn*, os modelos propostos da estatística *H* e o tratamento dos dados. A Análise Exploratória de Dados Espaciais, a análise gráfica das variáveis: média do Preço do Crédito e média do Custo

Médio, e os resultados da estimação do Índice de *Herfindahl-Hirschman* normalizado, e da estatística H de Panzar-Rosse são exibidos na quarta seção. Na quinta e última seção, são apresentadas as considerações finais.

2 REFERENCIAL TEÓRICO

Nesta seção, dividida em 6 subseções, são apresentados levantamentos objetivos dos principais tópicos abordados neste estudo. A primeira subseção tem o objetivo de apresentar as definições adotadas neste estudo para banco digital e intermediação financeira. Na segunda subseção expõe-se um breve relato sobre a Organização Industrial Empírica, focado na abordagem estrutural (da qual se disserta também sobre o índice de *Herfindahl-Hirschman* normalizado) e na abordagem não estrutural (da qual se expõe a estatística H de Panzar-Rosse). Ainda na segunda subseção, apresenta-se um resumo sobre os modelos de competição. A revisão da literatura que aborda e estima os níveis de concentração e de concorrência na indústria bancária do Brasil encontra-se na subseção 3. Por fim, na subseção 4, são apresentadas as evoluções normativas do Banco Central que corroboram com a hipótese de que este atua na direção de estimular a concorrência na indústria bancária e de proporcionar um ambiente propício e atrativo para que novas firmas entrem neste mercado.

2.1 DEFINIÇÕES

Segundo o BACEN, os modelos de negócio digitais baseiam-se na inovação tecnológica, nos protocolos de comunicação virtual – sem a necessidade de presença física –, na aplicação de inteligências artificiais, na terceirização de tecnologias e no armazenamento de dados.

Em face desses elementos caracterizadores, o atributo de banco digital pode ser considerado mais propriamente como uma medida de grau do que de gênero, uma vez que na atualidade todas as instituições financeiras apresentam, em menor ou maior grau, essas características. Um elemento adicional para identificação dos bancos digitais seria a extensão em que não apenas os processos e sistemas de relacionamento com os clientes se apresentam digitalizados, como também os sistemas internos de processamento e armazenamento de dados estão adaptados às novas tecnologias (o que normalmente envolve terceirização massiva, infraestruturas compartilhadas e uso de bases de dados em nuvem). Nesse plano, o processo de transformação vem sendo mais lento, especialmente para as instituições maiores, que possuem sistemas legados ou próprios abrangentes, personalizados e complexos, que exigiram elevados investimentos de amortização diferida. (BANCO CENTRAL DO BRASIL, 2019b, p. 147)

Quanto à origem dos bancos digitais, segundo BANCO CENTRAL DO BRASIL (2020), existem três tipos: “(i) bancos digitais nativos, que já são constituídos integralmente no modelo digital; (ii) bancos digitais transformados, que resultam da digitalização

dos bancos convencionais; e (iii) representações digitais dos bancos tradicionais, que são bancos ou unidades de negócio especificamente criados ou anunciados sob o modelo digital pelos conglomerados bancários convencionais.”(BANCO CENTRAL DO BRASIL, 2020, p. 166)

Em suma, a intermediação financeira consiste em captar recursos de agentes superavitários via taxa de captação e transferi-los para os agentes deficitários, mediante a cobrança de outras taxas. Entretanto, como pontua Santos (2018), a intermediação financeira não consiste apenas na transferência dos recursos captados com depósitos para o mercado de crédito. Em certo grau, os segmentos de depósito são independentes dos segmentos de crédito, uma vez que existem regulamentações específicas para as opções de depósitos em poupança, à vista e a prazo. Regulamentações que balizam tanto o poder de transferência quanto a livre alocação do recurso.

2.2 ORGANIZAÇÃO INDUSTRIAL EMPÍRICA

Indiscutivelmente uma grande parte das transações econômicas ocorre por meio dos mercados. Dadas as condições, fatores exógenos e endógenos dos mercados, as empresas tomam decisões estratégicas que afetam o bem-estar dos agentes econômicos. Tais decisões são o cerne das discussões para o campo da Organização Industrial (IO), de acordo com Belleflame e Peitz (2010). Destas, as definições mais importantes são a de empreender (ou não) e a de continuar ou encerrar suas atividades neste mercado. Para isto, as firmas observam o cenário preexistente, as forças que agem a favor e contra, as barreiras de entrada e saída, as perspectivas futuras e, principalmente, o nível de competição do setor.

Para se verificar o nível de competição de um mercado, uma das metodologias é analisar o poder de mercado das firmas. Belleflame e Peitz (2010) apontam que o “[...]poder de mercado pode ser definido como a habilidade de elevar os preços acima do nível da competição perfeita. A primeira maneira de mensurar o poder de mercado é olhando a diferença entre o preço e o custo marginal.” (BELLEFLAME; PEITZ, 2010, p. 34, tradução nossa).⁷

De acordo com Silveira (2017), do problema central da teoria da Organização Industrial – aferir a competição é avaliar o funcionamento do mercado – deu-se origem à abordagem estrutural. Nesta direção, Joe Bain e Edward Mason, da tradicional escola de Harvard, deram luz ao paradigma da estrutura-conduta-performance, cujo pressuposto é: a **estrutura** do mercado (quantidade de produtores, nível de diferenciação dos produtos e serviços, níveis dos custos, tamanho da demanda, barreiras de entrada, tamanho e tipo

⁷ “Market power can be defined as the ability to raise prices above the perfectly competitive level. The first way to measure market power is to look at the difference between price and marginal costs.” (BELLEFLAME; PEITZ, 2010, p. 34)

das firmas, presença de economias de escala) determina a **conduta** ou estratégia tanto dos produtores quanto dos consumidores (estratégia de preços, investimentos, *marketing*, inovação) que, por sua vez, resulta na **performance** da indústria (lucro econômico, criação de valor/valor agregado, razão preço/custo marginal, retorno aos acionistas).

Em síntese, o paradigma da estrutura-conduta-performance alega que, quanto maior o nível de concentração de um mercado, mais as firmas tendem a adotar práticas anti-competitivas e, assim, fixar preços acima do seu custo marginal, auferindo maiores lucros. Neste sentido, verifica-se que os trabalhos baseados na abordagem estrutural têm como foco o tamanho e o número de empresas. Logo, dentre as técnicas existentes, o índice de *Herfindahl–Hirschman* normalizado se destaca. Este índice é largamente utilizado na literatura e amplamente adotado por instituições de defesa da concorrência. No Brasil, o BACEN vem utilizado o IHH_n como instrumento acessório no cálculo do nível de concentração econômica do setor financeiro desde de 2017. O índice é definido como a soma das quotas de mercado ao quadrado para todas as n empresas ativas no mercado:

$$IHH_n = \sum_{i=1}^n (IF_i)^2, \quad \forall i = 1, \dots, n \quad (1)$$

em que IF representa a variável de análise da i -ésima instituição financeira e n o número de empresas.

Dada a utilização da soma da participação no mercado das firmas ao quadrado, o índice captura a importância das empresas maiores, atribuindo um peso maior para elas.

Das fragilidades dos estudos baseados na abordagem estrutural destacam-se duas: a ausência das relações de causalidade e o fato de a concentração ser dada como exógena quando, de fato, pode ser endógena, produto das diferenças produtivas das firmas daquela indústria (SILVEIRA, 2017). Assim, a essência do problema parece estar na complexidade em concluir o que os diferentes níveis de concentração indicam. Contudo, entidades como o *Federal Trade Commission* (FTC), o *U.S. Department of Justice* (USDOJ) e o BACEN publicam os valores que utilizam para balizar suas políticas baseados em suas compreensões de mercados concentrados. No quadro 1 observa-se as interpretações dos intervalos de valores adotadas pelo BACEN ao IHH_n .

Quadro 1: Interpretação do índice *Herfindahl–Hirschman* normalizado.

Nível de concentração	Valores de referência.
1) Baixa	entre 0 e 0,1000
2) Moderada	acima de 0,1000 até 0,1800
3) Elevada	acima de 0,1800 até 1

Fonte: Banco Central do Brasil, organizado pelo autor.

Nota: O quadro foi construído com dados coletados nas notas de rodapé do Relatório de Economia Bancária de 2019, página 123, do Banco Central do Brasil.

A partir da década de 1980 nasce a metodologia conhecida como Nova Organização Indústria Empírica (NEIO). Os novos estudos buscam compreender o comportamento dos agentes econômicos, dos consumidores e dos produtores através de modelos propostos utilizando econometria.

Neste cenário, onde o poder de mercado e a concentração de mercados são variáveis endógenas do modelo, e as decisões dos agentes econômicos são interdependentes, a teoria dos mercados contestáveis de Baumol, Panzar e Willig (1983) têm bastante destaque. Segundo ela, em um mercado perfeitamente contestável – informações perfeitas, acesso irrestrito à tecnologias de produção e livre oportunidade comercial –, mesmo com apenas uma firma, o custo marginal seria igual a receita marginal. Com isto, a firma deve auferir lucro econômico zero, visto que, sem barreiras de entrada e saída e com custos zero de entrada e saída, preços praticados acima do custo marginal induzirá outras firmas a entrar na indústria.

Utilizando técnicas de estática comparativa, Panzar e Rosse (1987) demonstraram que modelos simples de equilíbrio de mercados com produto único – competição monopolista, oligopólio e perfeitamente competitivos – geram restrições examináveis da equação de receita reduzida das firmas. Embora esta questão seja mais complexa ao lidar com firmas que ofertam múltiplos produtos e demandam múltiplos insumos, “[...] a hipótese de maximização do lucro do monopólio coloca restrições testáveis sobre os parâmetros estimados da equação de receita das empresas na forma reduzida.” (PANZAR; ROSSE, 1987, p. 443, tradução nossa).⁸

A estatística H corresponde à soma das elasticidades da receita com relação aos preços dos insumos de produção e é derivada por Panzar e Rosse (1987) da seguinte forma: Assume-se que y seja um vetor de variável de decisão que afeta a receita da firma e z um vetor de variável exógena que muda a função receita da firma, de tal jeito que a função receita seja $R = R(y, z)$. O vetor y não deve ser interpretado restritivamente como nível de produção, ele pode ser qualquer variável de decisão que faça parte da estrutura da função de receita da firma monopolista. O w é um vetor de m preços de fatores de produção que são exógenos à firma e o t é um vetor de variáveis exógenas que muda a função custo da firma. Também é assumido que os custos da firma dependem direta ou indiretamente de y , de tal forma que a função custo possa ser representada como $C = C(y, w, t)$. Deste jeito a função lucro (π) pode ser escrita como $\pi = R - C = \pi(y, z, w, t)$.⁹ Consequentemente, admite-se que a firma i maximiza o seu lucro no ponto em que a receita marginal (R'_i) é igual ao seu custo marginal (C'_i).

Logo, a estatística H representa a variação percentual da receita de equilíbrio derivada de uma oscilação percentual no preço dos m fatores de produção.

⁸ “The hypothesis of monopoly profit maximization does place testable restriction on the estimated parameters of firms’ reduce form revenue equations.” (PANZAR; ROSSE, 1987, p. 443)

⁹ Para a derivação completa da estatística H . Ver: (PANZAR; ROSSE, 1987).

$$\psi^* \equiv \sum \frac{w_i}{R^*} \left(\frac{\partial R^*}{\partial w_i} \right) \leq 0 \quad (2)$$

Onde, R^* : representa a receita marginal da empresa i ; w_i : representa o somatório dos insumos exógenos da empresa i .

Uma vez que a hipótese nula da estatística H é monopólio, se estabelece um resultado geral denominado:

Teorema 1 “A soma das elasticidades-preço dos fatores da equação de receita na forma reduzida de um monopolista deve ser não positiva.” (PANZAR; ROSSE, 1987, p. 445, tradução nossa) ¹⁰

Ademais, é importante salientar junto a este resultado as duas proposições centrais do modelo.

“**Proposição 1:** No equilíbrio simétrico de Chamberlin, a soma das elasticidades da receita da firma na forma reduzida, em relação ao preço dos fatores, é menor ou igual a um.” (PANZAR; ROSSE, 1987, p. 451, tradução nossa) ¹¹

“**Proposição 2:** Para firmas observadas em equilíbrio competitivo de longo prazo, a soma das elasticidades da receita da firma na forma reduzida, em relação ao preço dos fatores, é igual a um.” (PANZAR; ROSSE, 1987, p. 452) ¹²

Dado que o modelo Panzar-Rosse é baseado no somatório das elasticidades das receitas com relação aos insumos dos fatores de produção, a estatística H pelo método dos Mínimos Quadrados Ordinários pode ser estimada pela equação 3. (RONDON, 2011)

$$\ln R_i = \alpha + (\ln \sum w_i) \beta + \ln z_i \gamma + \varepsilon \quad (3)$$

Com o objetivo de facilitar a interpretação da estatística H – que é, em suma, o coeficiente β ¹³ resultantes da estimativa econométrica – Kadir et al. (2014), Silveira (2017) e Neto, Araújo e Ponce (2006), apresentam versões do quadro 2 em que o H menor ou igual a zero caracteriza um mercado ausente de pressões competitivas, o H entre zero e um indica competição monopolística. Já o H igual à um, designa um mercado perfeitamente competitivo.

¹⁰ “Theorem 1: The sum of the factor price elasticities of monopolist’s reduce form revenue equation must be nonpositive.” (PANZAR; ROSSE, 1987, p. 445)

¹¹ “Proposition 1: In symmetric Chamberlinian equilibrium, the sum of the elasticities of firm’s reduce form revenues with respect to factor prices is less than or equal to unity.” (PANZAR; ROSSE, 1987, p. 451)

¹² “Proposition 2: For firms observed in long-run competitive equilibrium, the sum of the elasticities of reduce form revenues with respect to factor prices equals unity.” (PANZAR; ROSSE, 1987, p. 452)

¹³ O β é o coeficiente dos w ’s vetores dos m preços dos fatores de produção que são exógenos à firma

Quadro 2: Interpretação da estatística H de Panzar-Rosse.

Valor H	Interpretação
$-\infty < H \leq 0$	Monopólio ou Oligopólio perfeitamente conivente.
$0 < H < 1$	Competição monopolística.
$H \geq 1$	Concorrência perfeita, monopólio natural ou mercado perfeitamente contestável.

Fonte: Adaptado de (KADIR et al., 2014), (SILVEIRA, 2017) e (NETO; ARAÚJO; PONCE, 2006).

A estatística H é teoricamente bem fundamentada e bem acolhida pela comunidade científica, porém, como mencionado por Divino e Silva (2017), sua aplicação requer cuidados especiais. O pressuposto básico do modelo é que a receita da firma monopolista ou oligopolista responde negativamente a incrementos nos preços dos fatores de produção. O modelo desenvolvido da estatística H foi aplicado a uma firma em tempo discreto, assim, como assinalado por Panzar e Rosse (1987), para aplicações empíricas apropriadas se faz necessário incorporar ao modelo variáveis capazes de captar detalhes institucionais do mercado em estudo. Entretanto, Bikker, Shaffers e Spierdijk (2012) apontam que o uso de variáveis de escala para controle do tamanho das instituições, como em Neto, Araújo e Ponce (2006), induzem a resultados viesados. Também é importante lembrar que na abordagem não estrutural, variáveis estruturais são endógenas ao modelo. Logo, o que se observa é o uso de uma variável ou *proxy* para a demanda individual das instituições.

Críticas mais fortes a estatística H são feitas por Bikker, Shaffers e Spierdijk (2012) e Spierdijk e Shaffer (2015). Estes autores apontam que em diversos modelos, com os resultados obtidos na estimação da estatística H , não é possível distinguir conluio e competição, uma vez que, em alguns resultados obtidos por eles, verificou-se $H < 0$ em mercados altamente competitivos. Observada as limitações do modelo da estatística H de Panzar-Rosse, verifica-se a necessidade de medidas complementares para validar as conclusões obtidas com as estimativas. Neste sentido, neste estudo utiliza-se o $IHHn$ e a análise gráfica das *proxies* para a média do Preço do Crédito e média do Custo Médio trimestral da indústria bancária.

2.2.1 Modelos de competição - resumo

Em um mercado de competição perfeita, as firmas são tomadoras de preço. Elas ofertam seus produtos ou serviços à preços correntes, ou seja, elas não têm poder de influenciar o preço que se paga por seus produtos ou serviços, independente da quantidade produzida. Usualmente é um mercado composto por muitos compradores e vendedores. Neste cenário o preço é igual ao custo marginal (CMg) – custo referente a produção de uma unidade adicional -. Já na concorrência imperfeita, as firmas são formadoras de

preços. Elas podem influenciar os preços de seus produtos, mas somente podem aumentar a quantidade vendida por meio da redução de seus preços. Na ótica dos consumidores, os produtos ou serviços ofertados por cada firma são diferenciáveis.

No mercado em que as firmas são formadoras de preços, são necessárias metodologias adicionais para descrever como são formados preços e as quantidades que sustentam a economia. A estrutura de mercado de concorrência imperfeita mais simples é o monopólio puro, que é um mercado no qual a firma não enfrenta concorrência. A curva da receita marginal (RMg) – receita extra que a firma ganha ao vender uma unidade adicional – da empresa monopolista está situada abaixo da curva de demanda, isto é, a receita marginal para esta firma é menor que o preço, uma vez que, para vender uma unidade adicional, a firma tem que baixar o preço de todas as unidades e não somente da unidade marginal. A relação entre a RMg e o Preço deriva de dois aspectos: da quantidade vendida e da inclinação da curva de demanda. Quanto mais horizontal for a curva de demanda, menos o monopolista terá que reduzir seu preço para acomodar a unidade adicional, de modo que a RMg está próxima do preço por unidade. Em contrapartida, quanto mais vertical for a curva de demanda, a venda da unidade adicional resultará em uma grande queda dos preços, de forma que a RMg seja menor que o preço. A inclinação da curva de custo médio (CMe) – custo total dividido pela quantidade produzida – é negativa, refletindo a hipótese de existência de economias de escala na indústria, de maneira que, quanto maior for a produção da firma, menor são os custos por unidade produzida. (KRUGMAN; OBSTFELD, 2001)

Na prática, a estrutura em indústrias caracterizadas por economias de escala interna são os oligopólios. O oligopólio é um mercado composto por apenas algumas firmas, não há livre entrada e a lucratividade do setor depende do modo como as empresas interagem. Um importante ponto básico da teoria do oligopólio é que o resultado do mercado sob competição imperfeita depende da variável preço ou quantidade. Isto é, cada firma determinará seu preço ou nível de produção, não apenas considerando a resposta do consumidor, mas também a estratégia do concorrente. Ou seja, as firmas são interdependentes e sua interação se dá através de um complexo jogo de estratégias, nos quais se destacam os modelos, Bertrand, Cournot e Stackelberg. Caso as interações tendam a ser mais cooperativas, as firmas podem cobrar preços muito acima do CMg, obtendo lucros maiores. Caso a concorrência seja mais agressiva, os preços tendem a se aproximar do CMg.

Krugman e Obstfeld (2001) destacam um caso especial de oligopólio denominado Concorrência Monopolística. Os mercados nesta estrutura possuem duas características principais: primeiro, as empresas competem vendendo produtos diferenciados, altamente substituíveis uns pelos outros, mas que não são substitutos perfeitos. Logo, a elasticidade cruzada da demanda das firmas é grande, mas finita. Segundo, não há barreiras consideráveis a entrada ou saída no mercado.

Assim como ocorre no monopólio, na competição monopolística as firmas se defron-

tam com a curvas de demandas descendentes, uma vez que a empresa é a única produtora de sua marca. A firma maximiza o lucro quando a receita marginal é igual ao custo marginal e o preço é maior que o custo médio, que ocorre quando o preço excede o CMg. No longo prazo, o lucro atrai novas firmas para o setor. A firma perde participação no mercado o que desloca sua curva de demanda para baixo e o preço torna-se igual ao CMe.

2.3 CONCENTRAÇÃO E CONCORRÊNCIA NA INDÚSTRIA BANCÁRIA BRASILEIRA

Observando os resultados do (IHH_n) para o segmento bancário comercial (b1) publicados pelo BACEN em 2019 e 2020 no Relatório de Economia Bancária, verifica-se uma constante redução nos níveis de concentração das variáveis Ativo Total, Depósito Total e Operações de Crédito no período entre 2017 a 2020. A mesma descendente nos níveis de concentração, a partir de 2017, é observada em Wolf (2020a), indicando que o segmento bancário brasileiro apresenta níveis moderados de concentração.

Contudo, no trabalho de Wolf (2020a), verifica-se que de 2010 até 2017 os níveis de concentração na indústria bancária nacional estavam em ascendência, chegando a superar o valor de 0,18 para operações de crédito, o que indica segundo os padrões do BACEN, elevado nível de concentração. Outros estudos, anteriores a 2017, que encontram elevados níveis de concentração para o segmento bancário comercial são Neto, Araújo e Ponce (2006) e Silveira (2017). Outro resultado interessante apresentado por Wolf (2020a) é quanto à concentração das cinco maiores instituições. O autor observa que a entrada das *fintechs* no mercado de produtos e serviços financeiros não alterou conclusivamente o comportamento do IHH das cinco maiores instituições no período entre 2008 e 2018.

A precursora dos estudos utilizando a estatística H de Panzar-Rosse para o setor bancário nacional foi Belaisch (2003). Com uma amostra de 49 bancos, no intervalo entre 1997 e 2000, a autora concluiu que os bancos brasileiros possuem um comportamento condizente com oligopólio. Outra importante conclusão deixada pela autora é que as estimativas mostram que para aquele período, os bancos brasileiros, apesar de serem lucrativos, eram muito pouco eficientes na intermediação financeira. A explicação encontrada é que o risco de crédito era alto e, graças a disponibilidade de títulos públicos com alto rendimento e livres de risco, os bancos foram capazes de arcar com a aversão. (BELAISCH, 2003).

Neto, Araújo e Ponce (2006) apresentam em seu artigo a oscilação semestral da estatística H e do (IHH_n) no intervalo temporal entre dezembro de 1995 e junho de 2004. Com pouca variação observada na estatística H para o período, os autores concluem que

o mercado bancário brasileiro da época atuava em regime de competição monopolística. O destaque do estudo está para a relação negativa entre concentração – (IHH_n) – e competitividade – estatística H – verificado nas estimações. Indicando que “[...] a maior concentração implica em menor competição no mercado.” (NETO; ARAÚJO; PONCE, 2006, p. 12). Os autores também relatam uma relação inversa entre o número de agências e a competição, apontando que, quanto maior a rede de agências, menor foi a competitividade observada no mercado.

Um dos artigos mais citados quando se estuda a estatística H para o mercado bancário brasileiro é o de Lucinda (2010). O estudo se destaca pela análise institucional em que o setor financeiro atravessava no período da amostra – primeiro trimestre de 2000 até o quarto trimestre de 2005 –, pelas observações às críticas feitas a estatística H e às proposições metodológicas. Lucinda (2010) aponta que a metodologia de Panzar-Rosse não é isenta de problemas e deduz que os resultados derivados da estatística H são encontrados somente sob a condição de longo prazo. Neste sentido o autor também destaca os apontamentos de Shaffer (1983), que afirma que, em equilíbrio de curto prazo, as estimativas da estatística H são reduzidas para uma única cauda, na qual valores positivos rejeitam qualquer forma de competição imperfeita e $H < 0$. Ou seja, são consistentes com múltiplas estruturas de mercados competitivos.

Lucinda (2010) também assinala as críticas de Bikker, Spierdijk e Finnie (2006), que discorrem sobre como o uso de variáveis de escala, na variável independente, tende a viesar a estatística H para cima. Entretanto, o autor afirma que a metodologia promove um bom começo para o estudo do grau de competição, uma vez que não requer a definição de mercados relevantes. Dentre o diferencial metodológico, se destacam a aplicação da estatística a diferentes amostras – Bancos Grandes, Bancos Médios, Bancos Pequenos, Bancos Comerciais, Bancos de Investimento e Todos os Bancos – e o uso da variável depósito como fator de produção necessário para o banco ofertar crédito e como um produto (serviço) do banco.

Os primeiros resultados apontam que, no longo prazo, o mercado bancário não aparenta estar em equilíbrio, mas também não aparenta características de conluio. Já os resultados baseados na metodologia de Gollop e Roberts (1979) aplicadas por Moreno, Martínez e Ruiz (2006) – em que “[...] a suposição de igualdade nos parâmetros de conduta entre as empresas e período de tempo é relaxada, e, também é adotada a visão diferenciada da estrutura insumo-produto do banco” (LUCINDA, 2010, p. 853, tradução livre)¹⁴ – para a indústria bancária espanhola, indicam que, para algumas firmas e em determinados períodos de tempo, é verificável um comportamento cooperativo.

Assim como Neto, Araújo e Ponce (2006), Rondon (2011) também faz estimações

¹⁴ “[...] the assumption of equality of conduct parameters between firms and time period is relaxed, and is also adopted a different view of input-output structure of the banking firm.” (LUCINDA, 2010, p. 853)

semestrais da estatística H , porém no intervalo que percorre o primeiro semestre de 1995 até o segundo semestre de 2008. O autor ressalva não ser possível observar uma tendência. Contudo, verifica-se que, mesmo dentro de um movimento oscilatório, e principalmente a partir de 2004, verifica-se uma propensão de elevação do indicador, o que sugere aumento no nível de competição. As hipóteses nula, $H_0 \leq 1$ e $H_0 \geq 0$ foram rejeitadas em todos os períodos, mesmo no nível de 1% de significância, caracterizando o mercado bancário do Brasil como concorrência monopolística.

Barbosa, Rocha e Salazar (2014) estendem as estimativas da estatística H para investigar o impacto das operações de vários produtos no grau de competição. O ensaio foi realizado utilizando os bancos de dados para instituições independentes e para conglomerados financeiros entre o primeiro trimestre de 2001 e o quarto trimestre de 2012. Como resultado, os autores “[...] encontram que, bancos que oferecem produtos clássicos (ex: empréstimos e cartões de crédito) e outros produtos (ex: serviços de corretagem, seguro e títulos de capitalização) têm poder de mercado substancialmente maior do que os bancos que oferecem apenas produtos clássicos.” (BARBOSA; ROCHA; SALAZAR, 2014, p. 2, tradução nossa)¹⁵. Assim, eles afirmam que os testes de poder de mercado que desconsideram o multiproduto bancário tendem a obter resultados subestimados.

Já Tabak, Gomes e Jr (2015) utilizam a estatística H para analisar como a exposição ao risco dos bancos é afetada por seu grau de poder de mercado. Os autores estimam o H semestralmente, como feito por Neto, Araújo e Ponce (2006), contudo, para o intervalo temporal compreendido entre junho de 2001 e junho de 2011 e utilizando a técnica de Panzar-Rosse associada a uma técnica não paramétrica denominada *local regression*. Os resultados implicam que o setor bancário do Brasil é caracterizado pela competição monopolística e está negativamente relacionado a sua exposição ao risco, independente das oscilações dos níveis de capital dos bancos.

Assim como Barbosa, Rocha e Salazar (2014), Cardoso, Azevedo e Barbosa (2016) também abordam a modelagem de Panzar-Rosse utilizando os relatórios financeiro-contábeis das instituições financeiras independentes (relatório 4010) e dos conglomerados financeiros e instituições independentes (relatório 4040). Para dar robustez aos resultados, também são utilizados os modelos de Bonne (2008) e Bresnahan-Lau (1982). Com a base de dados para o período de 1999 a 2010, os resultados despontam que os estudos que desconsideram os efeitos da coordenação entre as empresas do mesmo grupo econômico estão sujeitos a viés de superestimação do nível de concorrência estimado para a indústria bancária do Brasil.

Outro artigo que se sobressai pelas proposições metodológicas é o de Divino e Silva (2017). Os autores, além de abordarem as ponderações de Bikker, Shaffers e Spierdijk

¹⁵ “[...] find that banks that offer classic products (e.g., loans and credit cards) and other products (e.g., brokerage services, insurance and capitalization bonds) have substantially greater market power than banks that offer classic products only.” (BARBOSA; ROCHA; SALAZAR, 2014, p. 2)

(2012), também inovam ao destacar a importância do ajuste de sazonalidade nos dados e a normalização das despesas administrativas por unidades de crédito concedido. Os resultados obtidos no ensaio também são bem diferentes dos demais estudos direcionados a economia bancária do Brasil. Em um painel balanceado com 62 instituições financeiras compreendido entre julho de 1994 e fevereiro de 2012, Divino e Silva (2017) reportam encontrar $H < 0$, apontando para comportamento não competitivo das firmas na indústria bancária. Os autores também destacam que, “os 5 maiores bancos não são competitivos e seu poder de mercado é consideravelmente maior que as demais instituições.” (DIVINO; SILVA, 2017, p. 418, tradução nossa)¹⁶. Segundo eles, esta pode ser a explicação para a elevada lucratividade dos bancos, tendo em vista que as elevadas taxas praticadas reduzem a demanda por crédito. No estudo também é destacado que os bancos públicos possuem menor poder de mercado que as demais instituições, mas isso não se dá graças a um possível comportamento competitivo, mas sim, graças as interferências políticas.

Utilizando uma amostra compreendida entre 2011 a 2016, Silveira (2017) estima duas equações para verificar o poder de mercado das firmas na indústria bancária nacional. A primeira, utilizando a versão original de Panzar-Rosse na qual a Receita Total é a variável dependente. A segunda, utiliza o *return on asset* (ROA) como teste para rejeitar a hipótese de equilíbrio, pois, em equilíbrio competitivo de longo prazo, preço é igual ao custo marginal que é igual ao custo médio. Neste mercado, o lucro econômico é igual a zero, logo o H -ROA é igual a zero. (SILVEIRA, 2017, p. 36). Em adicional, o autor também acresce *dummies* em cada trimestre com o objetivo de avaliar os efeitos do tempo – sazonalidade – na amostra. Os resultados do painel estimados em MQO e FGLS *Feasible Generalized Least Squares* apontam para a não probabilidade do mercado atuar em regime competitivo e, em menor grau, em regime monopolista.

Como diferencial dos demais trabalhos até aqui pesquisados, Santos (2018) estima a estatística H para a indústria – composta por 48 bancos com atividades ininterruptas entre 2009 e 2016 – por grupo de atividade – Desenvolvimento, Investimento, Banco do Brasil, Caixa Econômica Federal e Bancos Múltiplos, Comerciais e de Câmbio (MCC) – e por instituições individuais. Os resultados evidenciam que os bancos de Desenvolvimento, Investimento e MCC aparentam atuar em concorrência monopolística. Contudo, os grupos Banco do Brasil e Caixa Econômica Federal resultaram em $H < 0$. Dos resultados individuais, destaca-se que a estatística H para os Bancos Bradesco, Itaú-Unibanco e Santander foram iguais a zero. O autor ressalta que “[...] os cinco maiores bancos de varejo promovem o *lock-in* da intermediação financeira, preferindo emprestar para o governo do que manter posições no varejo.” (SANTOS, 2018, p. 98, tradução nossa). Esse resultado reforça a suposição de Belaisch (2003) da preferência por títulos públicos no lugar do risco de crédito ao varejo.

¹⁶ “The five biggest banks are not competitive and their market power is considerably higher than that of the other institutions.” (DIVINO; SILVA, 2017, p. 418)

Em posse de uma amostra de instituições bancárias de 16 países da América Latina, Hordones e Sanvicente (2020) visam testar a hipótese do modelo Estrutura-Conduta-Performance de Bain (1951). A hipótese é “[...] que a concentração (estrutura) diminui a competição (conduta), ocasionando a maior lucratividade (medida de performance) do setor.” (HORDONES; SANVICENTE, 2020, p. 140). A estimação foi por MQO e o resultado mostra o Chile e Uruguai como os países com o maior H da amostra, 0,80. Para o Brasil, foi encontrado um $H = 0,49$, que foi o menor valor dentre os 16 países. Contudo, todas as economias da região caracterizam-se como competição monopolísticas.

2.4 EVOLUÇÕES NORMATIVAS DO BANCO CENTRAL DO BRASIL

É uma hipótese deste ensaio que o Banco Central do Brasil trabalha para promover um ambiente propício à criação de novas instituições e do desenvolvimento digital do sistema financeiro nacional e, com isto, aumentar a concorrência na indústria sem comprometer a estabilidade do sistema. No intuito de dar robustez a esta premissa, elencam-se algumas resoluções condizentes com esta argumentação:

Resolução nº 3.919, de 25 de novembro de 2010 $\Rightarrow$ “Altera e consolida as normas sobre cobrança de tarifas pela prestação de serviços por parte das instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil e dá outras providências.” (BRASIL. Banco Central do Brasil, 2013a, p. 1). Com esta medida, o BACEN padroniza os pacotes de serviços ofertados pelas instituições financeiras e veta a cobrança de serviços essenciais a pessoas físicas. Entende-se com isto que o Banco Central acirra a concorrência por depósitos, uma vez que, com pacotes de serviços homogêneos e gratuitos, os consumidores tendem a migrar (ou abrir) suas contas para instituições em que o acesso aos produtos bancários é mais próximo ou fácil. É importante ressaltar, também, que esta resolução prevê o escopo para a oferta da conta eletrônica sem tarifa.

Resolução nº 4.282, de 4 de novembro de 2013 $\Rightarrow$ “Estabelece as diretrizes que devem ser observadas na regulamentação, na vigilância e na supervisão das instituições de pagamento e dos arranjos de pagamento integrantes do Sistema de Pagamentos Brasileiro (SPB), de que trata a Lei nº 12.865, de 9 de outubro de 2013.” (BRASIL. Banco Central do Brasil, 2013b, p. 1). Nesta resolução o BACEN estabelece os parâmetros que devem ser observados na regulamentação, na vigilância e na supervisão das instituições de pagamento e das disposições de pagamento integrantes do Sistema de Pagamentos Brasileiro (SPB). Também são dadas as diretrizes para a concessão de autorização para constituição e funcionamento das Instituições de Pagamento – Código N4 no IFdata – e normatiza as contas de pagamento.

Resolução nº 4.480, de 25 de abril de 2016 $\Rightarrow$ “Dispõe sobre a abertura e o encer-

ramento de contas de depósitos por meio eletrônico e dá outras providências.” (BRASIL. Banco Central do Brasil, 2016, p. 1). Com esta resolução, o BACEN estabelece os requisitos a serem observados para que as instituições financeiras possam ofertar contas de depósito através de plataformas eletrônicas (internet) sem que seja necessário o deslocamento físico até uma agência. “§1º Consideram-se meios eletrônicos os instrumentos e os canais remotos utilizados para comunicação e troca de informações, sem contato presencial, entre clientes e as instituições referidas no caput.” (BRASIL. Banco Central do Brasil, 2016, p. 1). Entende-se esta como a resolução mais importante para a concepção dos bancos %100 digitais.

Resolução nº 4.630, de 25 de janeiro de 2018 $\Rightarrow$ “Altera a Resolução nº 4.480, de 25 de abril de 2016, que dispõe sobre a abertura e o encerramento de contas de depósitos por meio eletrônico.” (BRASIL. Banco Central do Brasil, 2018a, p. 1). Estende a resolução 4.480 – abertura e o encerramento de contas de depósitos por meio eletrônico – para os microempreendedores individuais.

Resolução nº 4.639, de 22 de fevereiro de 2018 $\Rightarrow$ “Altera a Resolução nº 3.402, de 6 de setembro de 2006, que dispõe sobre a prestação de serviços de pagamento de salários, aposentadorias e similares sem cobrança de tarifas.” (BRASIL. Banco Central do Brasil, 2018b, p. 1). Nesta resolução o BACEN assegura a portabilidade da conta de depósitos ou conta de pagamento pré-paga na própria instituição ou em outra instituição bancária.

Resolução nº 4.656, de 26 de abril de 2018 $\Rightarrow$ “Dispõe sobre a sociedade de crédito direto e a sociedade de empréstimo entre pessoas, disciplina a realização de operações de empréstimo e de financiamento entre pessoas por meio de plataforma eletrônica e estabelece os requisitos e os procedimentos para autorização para funcionamento, transferência de controle societário, reorganização societária e cancelamento da autorização dessas instituições.” (BRASIL. Banco Central do Brasil, 2018c, p. 1). Com esta resolução o BACEN especifica duas novas modalidades de instituição financeira a atuar no mercado brasileiro, a Sociedade de Crédito Direto (SCD) e a Sociedade de Empréstimos entre Pessoas (SEP), regulamentando as *fintechs*. A partir desta normativa, as *fintechs* passam a ter autonomia a atuarem sem a obrigatoriedade de constituir parceria com alguma instituição financeira tradicionais.

Resolução nº 4.697, de 27 de novembro de 2018 $\Rightarrow$ “Altera a Resolução nº 4.480, de 25 de abril de 2016, que dispõe sobre a abertura e o encerramento de contas de depósitos por meio eletrônico.” (BRASIL. Banco Central do Brasil, 2018d, p. 1). Estende a resolução 4.480 – abertura e o encerramento de contas de depósitos por meio eletrônico – para todas as pessoas jurídicas.

3 METODOLOGIA

Nesta seção são apresentados, os dados, as variáveis, os tratamentos realizados nas bases de dados para a elaboração da: Análise Exploratória de Dados Espaciais; Análise Gráfica; estimação do Índice de *Herfindahl-Hirschman* normalizado; e da estimação da Estatística *H* de Panzar-Rosse.

3.1 DADOS

No ensejo do objetivo específico – breve discussão sobre a distribuição das agências bancárias no território brasileiro – foram coletados dados nos *sites* do Banco Central do Brasil (Estatística Bancária por Municípios de Dezembro de 2019¹⁷) do Instituto Brasileiro de Geografia e Estatística (Densidade populacional estimada¹⁸ e o *shapefile* da malha municipal de 2020 das localidades e limites de 2017¹⁹).

Para realizar o objetivo principal da pesquisa – analisar o impacto do crescimento dos bancos digitais sobre a estrutura de mercado da indústria bancária brasileira sob a ótica da concentração, utilizando o Índice de *Herfindahl-Hirschman* normalizado, e da competição, empregando a metodologia da estatística *H* de Panzar-Rosse – os dados foram coletados no *site* do Banco Central do Brasil, os balancetes contábeis trimestrais dos Conglomerados Financeiros e Instituições Individuais no período compreendido entre o primeiro trimestre do ano de 2000 até o quarto trimestre do ano de 2019.²⁰ Os quatro relatórios – Resumo, Ativo, Passivo e Demonstração dos Resultados – foram dispostos em um painel de dados que totaliza 137.215 observações e 85 variáveis dos 2.810 códigos de instituições financeiras.²¹

Para trazer as variáveis a valores de 2019, foi extraído no *sítio* do *ipeadata*²² o Índice Nacional de Preços ao Consumidor Amplo (IPCA) calculado e fornecido pelo IBGE. Utilizando a equação 4, apresentada por De-Losso (2020), é criado o fator de inflação do IPCA acumulado de cada trimestre do intervalo da amostra.

$$I_{t-1} = \frac{I_t}{(1 + \pi_t)} \quad (4)$$

¹⁷ (BANCO CENTRAL DO BRASIL, 2019a)

¹⁸ (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2019a)

¹⁹ (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2019b)

²⁰ (BANCO CENTRAL DO BRASIL, 2019c)

²¹ Os dados extraídos estão dispostos na base de dados do IFdata, que, por sua vez, são originários da COSIF – Sistema de Consolidação Contábil das Instituições Financeiras.

²² (INSTITUTO DE PESQUISA ECONÔMICA APLICADA, 2019)

Onde, I_{t-1} : representa o fator de inflação; I_t : representa o preço de uma cesta de consumo no período t ; π_t : representa a taxa de inflação do período t .

Em consonância com a literatura pesquisada e, principalmente, com o Relatório da Economia Brasileira do BANCO CENTRAL DO BRASIL (2019b), neste estudo foram selecionadas somente as instituições do segmento bancário comercial, isto é, o tipo de consolidado bancário b1 ²³. Dado este primeiro filtro, o Painel é reduzido de 137.215 observações para 8.296 de 210 códigos de instituições bancárias. Dos 210 códigos, 30 firmas não preenchem os requisitos necessários – não possuem valores nos verbetes de depósitos ou não realizam operações de crédito no período do estudo – para o cálculo da estatística H de Panzar-Rosse.

Também é importante ressaltar que alguns conglomerados possuem códigos diferentes em períodos distintos. Se não é encontrada explicação, período sem atividade do código, fusão, venda ou alteração do contrato social, é utilizado o código do conglomerado ou instituição individual do último relatório disponível. Consequentemente, o painel de dados possui, para o ensaio, 6.746 observações de 159 instituições bancárias. Cabe enfatizar que, em comparação com o setor, a amostra de 159 firmas após o filtro para o consolidado bancário b1 representa 91,77% dos depósitos totais e 83,42% das operações de crédito em dezembro de 2019, o que demonstra a representatividade da amostra.

A tabela 1 a seguir mostra a representatividade das variáveis Depósito Total (DT) e do Saldo das Operações de Crédito (SOC) no quarto trimestre dos anos de 2017, 2018 e 2019 em relação a amostra sem nenhum filtro ou tratamento (137.215 observações).

²³ b1 – Banco Comercial, Banco Múltiplo com carteira comercial ou Caixa Econômica.

Tabela 1 – Representatividade das amostras, filtrada para a classificação b1 e filtrada para a classificação b1 com tratamento, em comparativo a indústria bancária brasileira nos anos de 2017, 2018 e 2019.

Observações	Data	DT	% indústria	% b1	SOC	% indústria	% b1
137.215	dez/17	2538.62	100%		3367.62	100%	
137.215	dez/18	2776.14	100%		3356.01	100%	
137.215	dez/19	2895.55	100%		3363.14	100%	
8.296	dez/17	2370.42	93.37%	100%	2815.85	83.62%	100%
8.296	dez/18	2585.82	93.14%	100%	2800.46	83.45%	100%
8.296	dez/19	2676.20	92.42%	100%	2819.95	83.85%	100%
6.746	dez/17	2349.82	92.56%	99.13%	2804.20	83.27%	99.59%
6.746	dez/18	2566.44	92.45%	99.25%	2785.45	83.00%	99.46%
6.746	dez/19	2657.16	91.77%	99.29%	2805.56	83.42%	99.49%

Fonte: Banco Central do Brasil, organizado pelo autor.

Nota: DT = Depósito Total; SOC = Saldo das Operações de Crédito. A amostra sem nenhum filtro contém 137.215 observações. A amostra filtrada para a classificação b1 contém 8.296 observações. A amostra filtrada para a classificação b1 após todos os tratamentos contém 6.746 observações. % indústria é o comparativo das duas amostras para os valores totais obtidos dos dados da indústria. % b1 é o comparativo da amostra após o filtro de classificação b1 em comparativo a amostra filtrada em b1 com tratamento dos dados.

Logo, verifica-se que, mesmo com uma perda de 1550 observações, a amostra com as 6746 observações representa mais do que 99% das firmas na classificação b1, tanto na comparação dos depósitos, quanto para o saldo das operações de crédito. E, como já mencionado, em comparação com a indústria financeira, a amostra também é extremamente representativa.

3.2 VARIÁVEIS

As estimativas do índice de *Herfindahl-Hirschman* normalizado são realizadas pela equação 1 ($IHH_n = \sum_{i=1}^n (IF_i)^2$) – sendo as variáveis compostas via metodologia exposta pelo BANCO CENTRAL DO BRASIL (2019b) – na qual, é utilizado o Ativo Ajustado, (para o nível de concentração do Ativo Total); o saldo das carteiras de crédito (para as Operações de Crédito); e o Depósito Total menos os Depósitos Interfinanceiros mais as Letras de Crédito Imobiliário, Letras de Crédito do Agronegócio e Letras Financeiras (para os Depósitos Totais). O cálculo é feito para cada uma das três variáveis individualmente em cada um dos oitenta trimestres que compõem a amostra do estudo.

As *proxies* para o gráfico das curvas de Custo médio e de Preço do crédito são elaboradas a partir da fórmula de preço do crédito expostas no BANCO CENTRAL DO BRASIL (2018, p. 163), em que o Preço do Crédito é igual a receita da intermediação financeira dividida pelo estoque ou saldo das operações de crédito ($P^C = RIF/SOC$).

Neste sentido, o SOC é a quantidade de crédito e o Custo Médio é deduzido da divisão dos Custos Totais pelo Saldo das Operações de Crédito ($CM = CT/SOC$).²⁴

Para estimar o nível de competição da indústria bancária, é utilizado um modelo simplificado que se aproxime do modelo original de Panzar e Rosse (1987). Para fim de decomposição, será utilizada apenas a variável Receita da Intermediação financeira como a variável dependente. Entretanto, são estimados três modelos que se diferenciam apenas pelas variáveis resposta – Receita da Intermediação Financeira (RIF), Receita Total (RT) e *Return of Asset* (ROA). Neste sentido, a variável Receita da Intermediação Financeira (RIF) é o vetor de decisão que afeta a receita total da firma, o y do modelo original. Depósito Total (DT) é a variável exógena com poder de alterar a função receita da firma, o z do modelo original. Logo a função receita é:

$$R = R(y, z) = R(RIF, DT) \quad (5)$$

w é o vetor das m variáveis preço dos insumos de produção e foi composto pelo somatório das Despesas Administrativas (DA) do banco i no período t , Despesas de Pessoal (DP) do banco i no período t e Despesas de Captação (DC) do banco i no período t .

$$w_{it} = DA_{it} + DP_{it} + DC_{it} \quad (6)$$

A variável de controle CRD²⁵, o t do modelo original, é a variável exógena com poder de alterar a função custo total da firma. O CRD é uma *proxy* para o risco de exposição decorrente da intermediação financeira que influencia tanto a oferta quanto a demanda por fatores de produção. Enfim, o CRD é a variável que captura os detalhes institucionais, como sugerido por Panzar e Rosse (1987) para se obter testes estatísticos mais apurados.²⁶

Desta forma, a função custo fica:

$$C = C(y, w, t) = C(RIF, w, CRD) \quad (7)$$

e a função lucro é escrita como:

$$\pi(RIF, DT, w, CRD) = R(RIF, DT) - C(RIF, w, CRD) \quad (8)$$

Ademais, é adicionado ao painel uma coluna com variável *dummy* Tipo Bancário (TB)

²⁴ A composição dos custos totais encontra-se no quadro 3.

²⁵ O CRD é calculado seguindo Lucinda (2010).

²⁶ A estatística H também é estimada utilizando a Receita Total (RT) e o *Return of Asset* (ROA) como vetores de decisão.

que visa captar o diferencial de o banco ser (e, de quando passou a ser) digital. A firma recebe zero se não é (e, enquanto não é) digital, e um se é (e, a partir do momento que se tornou) banco digital. Os bancos foram classificados como digitais após verificação nos dados – alteração do contrato social – e nos *sites* das instituições. Das 8 instituições que atenderam os requisitos – ser b1, possuir valores nas variáveis mais importantes, demandar depósitos e ofertar créditos – somente uma é digital desde a concepção. As demais são resultado de fusões, aquisições e do deslocamento produtivo de instituições e conglomerados para a prestação de serviços bancários digitalizados. É importante ressaltar que bancos tradicionais também oferecem serviços no padrão dos bancos digitais, contudo, conforme a definição anteriormente apresentada, estes não são assim classificados.

O quadro 3 exibe as variáveis que são utilizadas neste estudo. Na primeira coluna – Modelo – são destacadas as três estimativas principais – $IHHn$, H de Panzar-Rosse e as Curvas analíticas. Na segunda coluna são apresentadas as variáveis e suas abreviações. Na terceira são apresentadas a composição de cada variável e, por fim, na quarta coluna são apontados os relatórios do IFdata no qual são encontrados os dados que compõe as variáveis.

Quadro 3: Variáveis utilizadas no estudo.

Modelo	Variáveis	Composição das Variáveis	Relatório no IFdata
IHHn	Ativo Total Ajustado (ATA)		Ativo
IHHn	Depósito Total (DTT)	Depósito Total - Depósitos Interfinanceiros + Letras de crédito Imobiliário + Letras de crédito do Agromercado + Letras Financeiras	Passivo
IHHn	Saldo das Operações de Crédito (SOC)	Operações de Crédito + Arrendamento Mercantil a Receber	Ativo
Estatística H	Receita da Intermediação Financeira (RIF)		Demonstração dos Resultados
Estatística H	Receita Total (RT)	Receita da Intermediação Financeira + Rendas de Prestação de Serviços + Renda de Tarifas Bancárias + Outras Receitas Operacionais.	Demonstração dos Resultados
Estatística H	Return of Asset (ROA)	Lucro Líquido / Ativo Total	Resumo
Estatística H	Depósito Total (DT)		Ativo
Estatística H	Insumos Exógenos (w)	Despesas de Pessoal + Despesas Administrativas + Despesas de Captação	Demonstração dos Resultados
Estatística H	Risco da Intermediação Financeira (CRD)	(Operações de Crédito + Arrendamento Mercantil a Receber)/(Depósito Total + Obrigações por Empréstimos e Repasses)	Demonstração dos Resultados e Ativo
Estatística H	<i>Dummy</i> Tipo Bancário (TB)		
Curva analítica	Preço do Crédito (PC)	Receita da Intermediação financeira / (Operações de Crédito + Arrendamento Mercantil a Receber)	Demonstração dos Resultados e Ativo
Curva analítica	Custo Médio (CM)	Insumos Exógenos (w) + Despesa de Intermediação Financeira + Despesa Tributárias + Outras Despesas Operacionais / (Operações de Crédito + Arrendamento Mercantil a Receber)	Demonstração dos Resultados e Ativo
Curva analítica	Número de Agências (NAG)		Resumo

Fonte: Banco Central do Brasil, organizado pelo autor.

3.3 MODELO EMPÍRICO PROPOSTO DA ESTATÍSTICA H DE PANZAR-ROSSE.

A estratégia empírica para estimar a estatística H de Panzar-Rosse é construir um painel de dados desbalanceado – não há observações de todos os indivíduos durante todo o intervalo temporal – com os dados percorrendo o intervalo temporal que parte do primeiro trimestre de 2000 até o quarto trimestre de 2019. Em contraste com dados em *cross-section*, em que vários indivíduos são observados em um ponto no tempo, os dados em painel, ou dados longitudinais, associam dados de diferentes indivíduos para períodos distintos. Isso é denotado por: $(X_{it}, Y_{it}), i = 1, \dots, n$ e $t = 1, \dots, T$, em que o index i faz referência ao indivíduo (Instituição Bancária) e o t ao período.

Segundo Gujarati e Porter (2011) os dados em painel se destacam pois: i) levam em consideração a heterogeneidade explicitamente, permitindo variáveis específicas ao indivíduo; ii) tendem a reportar estimativas mais eficientes e menos colineares entre variáveis; iii) podem minimizar o viés que pode resultar do trabalho com agregado de indivíduos e; iv) possibilita detectar e medir com maior qualidade os efeitos que não podem ser observados em um corte transversal.

Dado a decomposição das variáveis, obtém-se a equação da receita na forma reduzida em que $R(RIF, DT) = C(RIF, w, CRD)$. Desta forma, a função lucro pode ser escrita da seguinte forma: $\pi(RIF, DT, w, CRD) = R(RIF, DT) - C(RIF, w, CRD)$.

Logo, as especificações empíricas gerais dos modelos propostos são:

$$RIF_{it} = \alpha + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (9)$$

$$RT_{it} = \alpha + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (10)$$

$$ROA_{it} = \alpha + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (11)$$

Onde: o coeficiente β é a estatística H ²⁷, α refere-se ao parâmetro de intercepto, γ_1 , γ_2 e γ_3 são coeficientes do modelo, RIF_{it} é a Receita da Intermediação financeira da empresa i no período t , RT_{it} é a Receita Total da empresa i no período t , ROA_{it} é o *Return of Asset* da empresa i no período t , DT_{it} é o depósito total da empresa i no período t , w_{it} é o somatório dos insumos exógenos – Despesa de Pessoal da empresa i no período t + Despesas Administrativas da empresa i no período t + Despesas de Captação da empresa i no período t –, CRD_{it} é o Risco da Intermediação Financeira da empresa i no período t , TB_{it} é a variável *dummy* para o Tipo Bancário da empresa i no período t e

²⁷ Se $H \leq 0$: Monopólio ou Oligopólio perfeitamente conivente. Se H entre 0 e 1: Competição monopolística. Se $H \geq 1$: Concorrência perfeita, monopólio natural ou mercado perfeitamente contestável.

ϵ_{it} é o erro aleatório da empresa i no período t com média zero e variância constante não autocorrelacionados.

Neste estudo são consideradas as três possibilidades básicas de estimação de dados em painel: modelo (*Pooled*), modelo de Efeitos Fixos e modelo de Efeitos Aleatórios ou modelo de Componentes dos Erros. Segundo Greene (2000), exposto por Justus e Kassouf (2007), a estrutura básica do modelo de regressão de dados longitudinais, em notação matricial, é dada por:

$$y_{it} = x'_{it}\beta + z'_i\alpha + \epsilon_{it} \quad (12)$$

em que y_{it} é a variável dependente, x'_{it} é a variável independente, β é o coeficiente a ser estimado, ϵ_{it} é o termo do erro, α é a constante e z'_i é o conjunto de variáveis específicas de cada indivíduo, as quais podem ser ou não ser observados em todos as unidades de corte.

Adaptando a equação 12 na 9 obtem-se:

$$RIF_{it} = z'_i\alpha + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (13)$$

assim, RIF_{it} é a variável dependente, DT, w, CRD e TB são as variáveis independentes, β e γ são os coeficientes a serem estimados, α é a constante e z'_i é o conjunto de variáveis específicas de cada indivíduo, as quais podem ser ou não ser observados em todos as unidades de corte e ϵ_{it} é o termo do erro que supõe-se ser independente, identicamente distribuído com média zero e variância constante ($\epsilon_{it} \sim iid(0, \sigma_\epsilon^2)$).

O modelo *Pooled* – é, em suma, o modelo de mínimos quadrados ordinários (MQO) – assume que a heterogeneidade não-observada é igual para todos os indivíduos e, sendo assim, é absorvida pelo intercepto do modelo. A regressão através do modelo *pooled* fornece estimativas consistentes e eficientes se o z_i for observado para todos os indivíduos. Logo, sem heterogeneidade não observada, o $z_i\alpha$ é constante. Portanto, a equação 13 pode ser reescrita.

$$RIF_{it} = \alpha + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (14)$$

Entretanto, se é constatado a presença de heterogeneidade não observada em z_i , o β estimado será tendencioso e ineficiente em virtude da omissão da variável. Sendo assim, as análises devem ser baseadas nos modelos que permitem o controle da heterogeneidade não observada. Quando a heterogeneidade não observada é correlacionada com as variáveis dependentes (DT_{it} , w_{it} , CRD_{it} e TB_{it}) o modelo de Efeitos Fixos é preferível. Neste contexto, é pressuposto que $\alpha_i = z'_i\alpha$ seja um termo constante no tempo. Deste modo,

α_i é um parâmetro que necessita ser estimado, a variável independente e as variáveis dependentes recebem as n observações da i -ésima instituição financeira e ϵ_i recebe os n valores do termo do erro. Assim, o intercepto pode diferir entre os indivíduos, haja vista a heterogeneidade entre cada unidade de corte transversal. Deste modo, o modelo pode ser reescrito como:

$$RIF_{it} = \alpha_i + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (15)$$

onde, a $E(\epsilon_{it}|\alpha_i, x_{it}) = 0$, em que x_{it} é o conjunto das variáveis dependentes.

Por fim, se é constatado a presença de heterogeneidade não observada, mas ela não está correlacionada com as variáveis exógenas, opta-se pelo modelo de Efeitos Aleatórios. Neste modelo, supõe-se que o intercepto de uma unidade de corte transversal seja extraído aleatoriamente de uma população maior com um valor médio constante. O intercepto de cada indivíduo é expresso como um desvio desse valor médio constante. Assim sendo, o único termo constante é a esperança da heterogeneidade não observada ($E[z'_i \alpha]$). Portanto, a heterogeneidade não observada aleatória do i -ésimo indivíduo (v_i) é dada por: $v_i = (z'_i \alpha - E[z'_i \alpha])^{28}$. Reformula-se o modelo para:

$$RIF_{it} = \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + E[z'_i \alpha] + (z'_i \alpha - E[z'_i \alpha]) + \epsilon_{it} \quad (16)$$

que pode ser escrito como:

$$RIF_{it} = \alpha + \gamma_1 DT_{it} + \beta w_{it} + \gamma_2 CRD_{it} + \gamma_3 TB_{it} + v_i + \epsilon_{it} \quad (17)$$

com o termo do erro assumindo que a $E(\alpha_i^* + \epsilon_{it}) = 0$, em que α_i^* é uma variável aleatória, com média 0 e variância σ_α^2 .

3.4 ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS

Na tabela 2 é apresentada a estatística descritiva das variáveis para os modelos da estatística H de Panzar-Rosse – variáveis dependentes (Receita da Intermediação Financeira (RIF), Receita Total (RT) e *Return on Assets* (ROA) e independentes Depósito Total (DT), Insumos Exógenos (w), Risco da Intermediação Financeira (CRD) e *dummy* Tipo Bancário (TB) –, do $IHHn$ – Depósito Total (DTT), Saldo das Operações de Crédito (SOC) e Ativo Total Ajustado (ATA) –, e também, as duas *proxies*, preço do crédito

²⁸ Ver: Greene (2000), Gujarati e Porter (2011), Justus e Kassouf (2007) e Croissant e Millo (2008).

e custo médio²⁹, criadas para verificar o comportamento dos custos médios e dos preços ao longo do tempo. E assim, validar o uso da estatística H , haja vista as críticas de Spierdijk e Shaffer (2015) quanto aos resultados obtidos com o modelo.

Os valores apresentados na tabela 2 estão corrigidos pelo IPCA a preços acumulados do quarto trimestre de 2019. Observa-se que a variável CRD “explode” e reporta valores de máximo e mínimo tendendo ao infinito. Este comportamento ocorre dado a presença de *outliers* com valor zero na variável SOC. A variável RIF também apresenta *outliers* com valores negativos e zero, o que também faz as variáveis PC e CM reportarem *outliers*. Parte do comportamento atípico da variável CRD e RIF são corrigidos na sazonalização, no entanto para a transformação logarítmica e para a regressão, faz-se necessário eliminar ou tratar os dados.³⁰

As variáveis utilizadas para a estimação do $IHHn$ também apresentam alguns problemas. Contudo, uma vez que o cálculo é feito em tempo discreto, é possível eliminar as ocorrências do modelo, sem gerar reduções significativas na qualidade das estimativas. Verifica-se que a composição da variável DTT gera 80 números faltantes, que representa a diminuição expressiva no número de observações.

²⁹ As variáveis *proxy* do preço do crédito (PC) e do custo médio (CM) são estimadas com base na metodologia proposta no (BANCO CENTRAL DO BRASIL, 2018, p. 163).

³⁰ Uma série de testes são empregados para verificar o menor impacto nas séries temporais e nas estatísticas. Os resultados mostram que, no intuito de se obter um painel representativo, se faz necessário analisar caso a caso. Para algumas firmas, a presença de dados faltantes e dados comprometedores é muito grande e espalhada em diversos momentos. Nestes casos, faz-se necessário eliminar todas as observações da empresa. Em outros casos, os dados erráticos estão presentes e concentrados em um determinado intervalo temporal. Para estes, a solução proposta é ficar com o maior espaço temporal de dados utilizáveis. E, em raros casos, é verificado um ou dois *outliers* em todos os dados da empresa. Nestes casos em particular, verifica-se o tamanho da firma com relação a indústria, o impacto da ruptura da série temporal e o comportamento das médias e mediana ao aplicar o tratamento do *outlier*. É aplicada a estratégia com menor impacto nas estimativas.

Tabela 2 – Estatística descritiva das variáveis

Variável	N	Média	Desvio Padrão	Mínima	1º Quartil	3º Quartil	Máxima
RIF	6,746	3,048,601.000	10,269,914.000	−190,462.200	51,682.600	1,006,450.000	111,673,270.000
RT	6,746	3,938,059.000	13,430,603.000	183.238	66,466.820	1,283,202.000	148,335,383.000
ROA	6,746	0.008	0.058	−1.685	0.002	0.018	0.385
DT	6,746	22,168,413.000	79,696,105.000	0.000	236,159.300	5,511,193.000	704,425,892.000
w	6,746	2,247,913.000	7,639,458.000	−147,268.600	41,129.980	703,153.000	89,621,184.000
CRD	6,746	∞		0.000	0.823	2.074	∞
TB	6,746	0.008	0.087	0	0	0	1
DTT	6,666	23,811,243.000	90,047,264.000	0.000	228,143.400	5,289,758.000	760,018,182.000
SOC	6,746	23,004,236.000	89,152,011.000	0.000	252,458.900	6,511,983.000	812,113,751.000
ATA	6,746	63,701,969.000	222,286,070.000	21,642.280	844,215.800	19,236,943.000	1,723,632,154.000
PC	6,746	∞		−0.636	0.170	0.632	∞
CM	6,746	∞		−0.647	0.199	0.744	∞

Fonte: Banco Central do Brasil, organizado pelo autor.

Nota: RIF – Receita da Intermediação Financeira; RT – Receita Total; ROA – *Return on Assets*; DT – Depósito Total; w – Insumos Exógenos (Despesas de Pessoal + Despesas Administrativas + Despesas de Captação); CRD – Risco da Intermediação Financeira; TB – Tipo Bancário; DTT – Depósitos Totais (DT – Depósitos da Intermediação Financeira + Letras de Crédito Imobiliário + Letras de Crédito do Agronegócio + Letras Financeiras); SOC – Saldo das Operações de Crédito; ATA – Ativo Total Ajustado; PC – Preço do Crédito e CM – Custo Médio.

3.5 TRATAMENTO DOS DADOS

No ensejo de cumprir os objetivos propostos, alguns tratamentos no banco de dados se fazem necessários. Como já mencionado, no intuito de reduzir a heterogeneidade dos dados, usa-se a classificação b1 para segmento bancário comercial. Ademais, as variáveis são levadas ao valor de 2019 através da multiplicação pelo fator de inflação, para que seja eliminado os efeitos da inflação nos dados. É feita análise individual das séries temporais das firmas para sanar o problema com dados faltantes, *outliers* e, também dos problemas causados nas séries em virtude das fusões e aquisições. Nas seguintes subseções, são apresentadas a metodologia para dessazonalizar as séries temporais, para verificar a presença de raiz unitária, os tratamentos empregados, a estatística descritiva das variáveis tratadas e a matriz de correlação.

3.5.1 Sazonalidade

De forma clássica, uma série temporal pode ser decomposta em quatro componentes: tendência, sazonalidade, ciclo e erro. A sazonalidade é gerada por movimentos oscilantes de mesma periodicidade que ocorrem dentro de um intervalo anual, e sua ocorrência pode levar a conclusões equivocadas a respeito da série temporal. Ferreira e Mattos (2016) apontam que o ajuste sazonal ou dessazonalização – remoção da sazonalidade – pode ser alcançando por diversas metodologias, como, por exemplo, a utilização de *dummies* sazonais, modelos de suavização exponencial ou programas específicos como o X-13ARIMA-SEATS³¹.

Neste trabalho opta-se pela utilização do programa X13 ARIMA-SEATS no modo automático. Segundo Ferreira, Jr e Mattos (2015), a execução do programa automaticamente pode originar um ajuste sazonal de boa qualidade, pois verifica a ocorrência de *outliers*, a existência de efeitos de calendário, a necessidade de transformação log e a ordem do modelo ARIMA.

Para fins ilustrativos, são exibidos os resultados dos testes de sazonalidade das variáveis Receita da Intermediação Financeira (RIF) e Depósito Total (DT) do banco código 10045 – selecionado aleatoriamente. A mesma análise é realizada para 8 variáveis³²

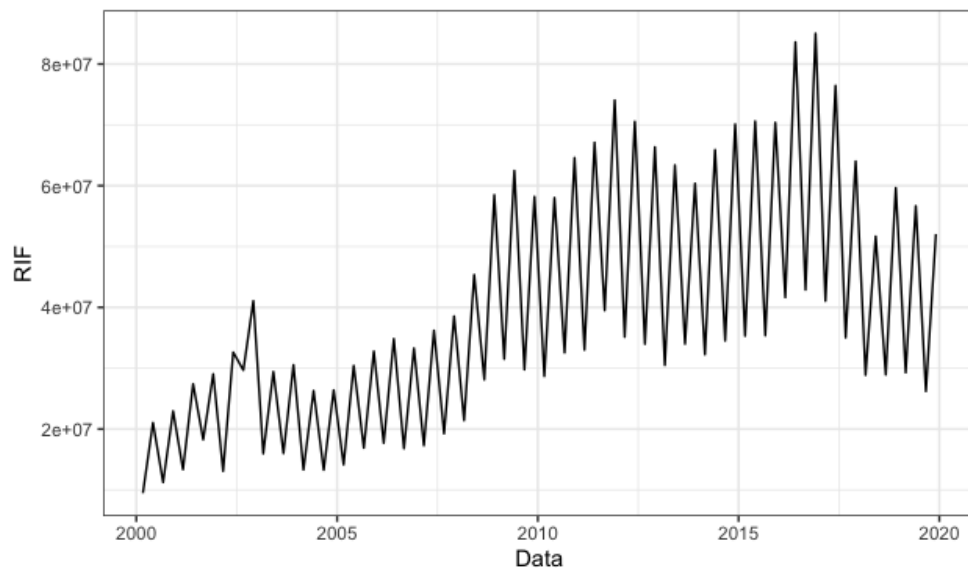
³¹ O X-13ARIMA-SEATS é um programa de ajuste sazonal desenvolvido em julho de 2012 pelos *U.S. Census Bureau* e o Banco da Espanha. O programa é uma evolução do X11 ARIMA de Estella Dagum que foi atualizado para X12 ARIMA pelo *U.S. Census Bureau* e que em 2012 foi novamente atualizado e integrado ao TRAMO/SEATS de Gomes e Maravall (1997) do Banco da Espanha. Disponível em: <https://www.census.gov/data/software/x13as.html>

³² RIF – Receita da Intermediação Financeira; RT – Receita Total; ROA – *Return of Asset*; DT – Depósito Total; w – $\sum$ Despesas de Pessoal, Despesas Administrativas e Despesas de Captação; CRD – *proxy* para o Risco da Intermediação Financeira; PC – Preço do Crédito e CM – Custo Médio.

de todas as instituições com mais de 20 observações. Mostrado no apêndice A, quadro 14. Conforme Ferreira, Jr e Mattos (2015), para executar a dessazonalização das séries temporais são seguidos 4 passos: i) Análise Gráfica; ii) Execução do programa; iii) Avaliação do ajuste sazonal; iv) Correção do ajuste sazonal (se necessário).

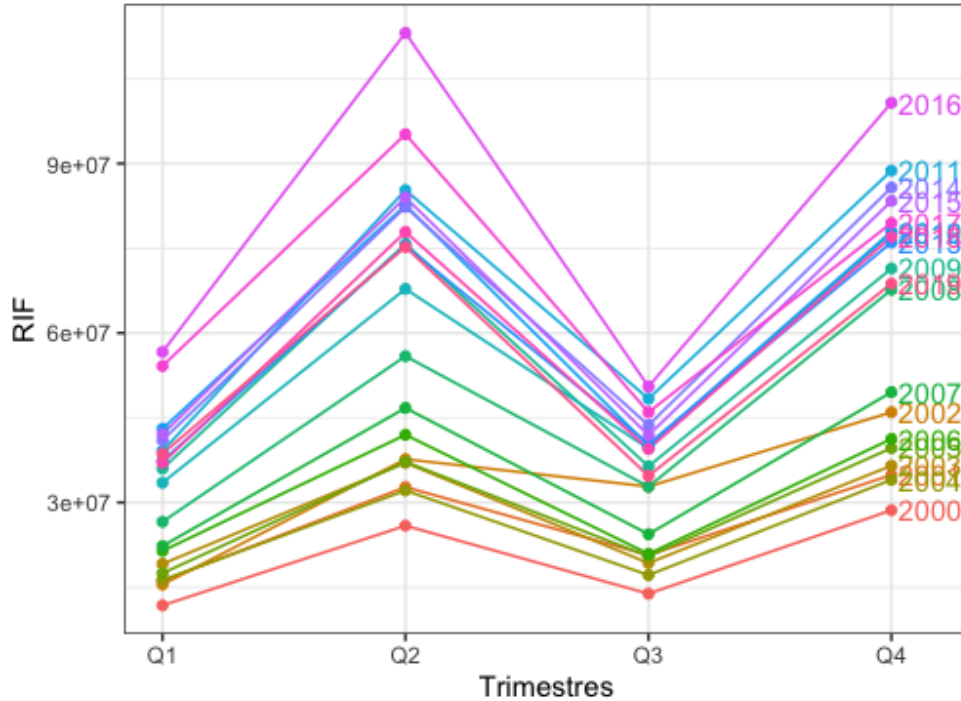
O gráfico da figura 3 representa o comportamento histórico da série temporal da variável Receita da Intermediação Financeira (RIF) ao longo dos 20 anos de observação. No eixo Y estão representados os valores da RIF variando entre R\$9.450.621,00 a R\$85.027.250,00. Já a figura 4 reporta o comportamento trimestral da variável destacando ano a ano as oscilações da RIF.

Figura 3 – Comportamento temporal da variável Receita da Intermediação Financeira da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.



Fonte: Banco Central do Brasil, organizado pelo autor.

Figura 4 – Oscilação anual da variável Receita da Intermediação Financeira da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.



Fonte: Banco Central do Brasil, organizado pelo autor.

Nota: Q1 – primeiro trimestre, Q2 – segundo trimestre, Q3 – terceiro trimestre e Q4 – quarto trimestre.

Por meio da análise gráfica é plausível supor que a variável Receita da Intermediação Financeira (RIF) possui característica sazonal, uma vez que ao passar dos anos os movimentos oscilatórios nos trimestres são visualmente similares. Entretanto, os movimentos não parecem tão constantes e é possível verificar alguns picos indicando que talvez sejam necessárias transformações para estabilizar a série.

Após visualizar os gráficos da variável, o próximo passo é executar o programa para testar e avaliar o ajuste sazonal. Segundo Ferreira e Mattos (2016), uma forma de se fazer isso é aplicando a estatística QS³³ na série temporal. A estatística QS é computada examinando a função de autocorrelação nas defasagens sazonais. Além de ser aplicado na série original e na série com ajuste sazonal, o teste também é aplicado ao componente irregular e aos resíduos do modelo ARIMA. A hipótese nula do teste pressupõe a inexistência de sazonalidade.

³³ Seja F a frequência sazonal anual, então $f = 4$ para dados em trimestre de cada ano e n é o tamanho da amostra. Ademais $\gamma(g) = E[y_{n+g}y_n] \sim E^2[y_n]$ representa a autocovariância e $\rho(g) = \gamma(g)/\gamma(0)$ é a correlação de Y . Logo a estatística QS pode ser escrita como:

$$QS = n(n+2) \left(\frac{\hat{\rho}^2(f)}{n-f} + \frac{[\max\{0, \hat{\rho}^2(2f)\}]^2}{n-2f} \right)$$

Note que $H_0 : \gamma(m) \leq 0$ para $m \in \{f, 2f\}$. Isso leva a uma restrição no domínio de $\gamma(m)$ de tal forma que $\hat{\gamma}(m) \leq 0$, assim $QS = 0$. Ver: (CORNWALL; CHEN, 2019, p. 6)

Tabela 3 – Estatística QS da variável Receita da Intermediação Financeira da instituição código 10045.

Ajuste	qs	p-valor
qsori	148.43926	0.00000
qsorievadj	148.21257	0.00000
qrsrd	0.38748	0.82387
qssadj	0.62501	0.73161
qssadjevadj	0.01525	0.99240
qsirr	0.00118	0.99941
qsirrevadj	9.84956	0.00726

Fonte: Resultado da pesquisa.

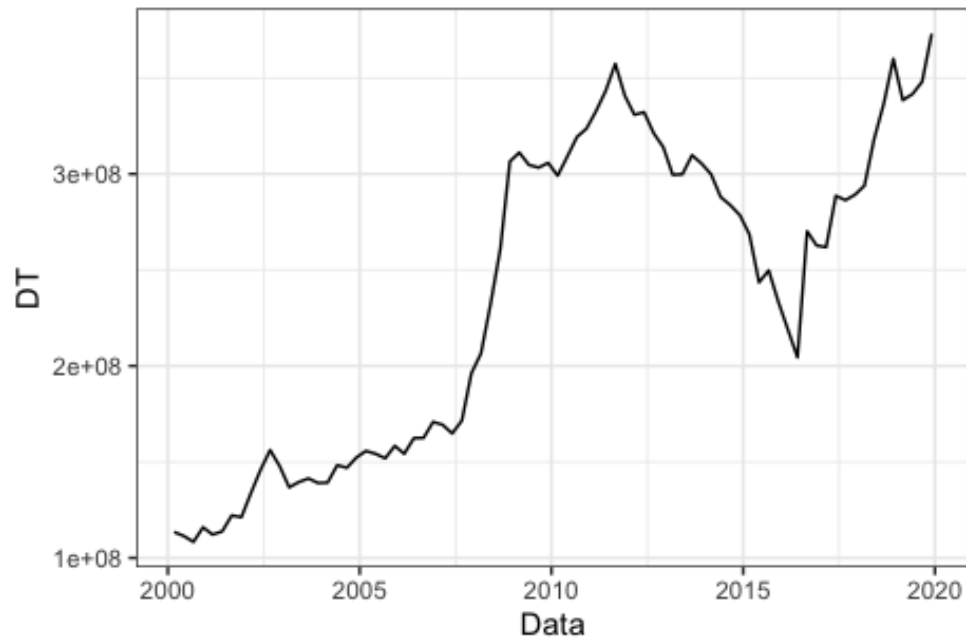
Nota: qsori – série original; qsorievadj – série original corrigida por valores extremos; qrsrd – resíduos do modelo ARIMA; qssadj – série temporal com ajuste sazonal; qssadjevadj – série temporal com ajuste sazonal corrigida por valores extremos; qsirr – componente irregular; qsirrevadj – componente irregular corrigido por valores extremos.

Como se pode observar na tabela 3, o *p-value*, tanto para a série original (qsori) como para a série original corrigida para valores extremos (qsorievadj), é menor que 5%, logo há forte evidência de que a série original (qsori) têm comportamento sazonal. Portanto, a hipótese nula do teste é rejeitada. Da mesma forma, observa-se o *p-value* superior a 5% na série temporal com ajuste sazonal (qssadj), indicando que o programa cumpriu seu papel em remover o comportamento sazonal da série temporal.

O diagnóstico da série temporal ajustada, da variável RIF, é apresentado em sua totalidade no apêndice B. Verifica-se que foi detectado um *outlier* no primeiro trimestre de 2002 (A02002.3) ao nível de significância de 1%. O modelo ARIMA estimado automaticamente é da ordem (3 1 1)(0 1 1), com todos os parâmetros significativos com 99% de confiança. O teste de autocorrelação Ljung-Box aponta que não há evidências de autocorrelação residual para o modelo ARIMA estimado. Apesar de não ser uma característica necessária nos diagnósticos de modelos ARIMA, o teste de normalidade Shapiro-Wilk indica não normalidade.

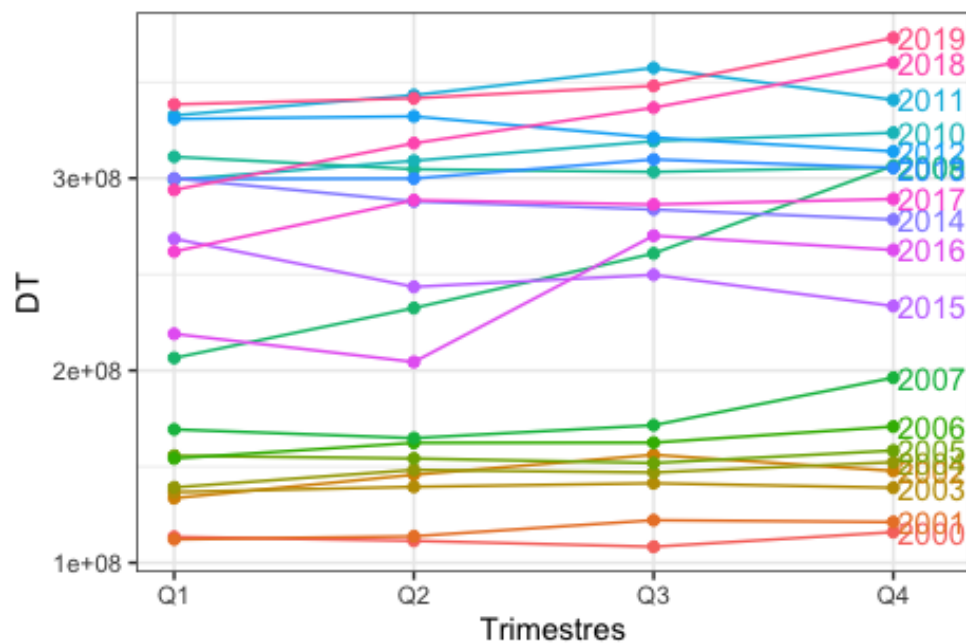
A análise gráfica (figuras 5 e 6) da série temporal da variável Depósito Total (DT)), também do banco 10045, já mostra um comportamento oposto ao verificado nas figuras 3 e 4 referentes a variável RIF. O gráfico da figura 5 esboça uma tendência ascendente sem movimentos oscilatórios constantes e a figura 6 tende a confirmar que os Depósitos Bancários não parecem ter variações anuais constantes, o que, por sua vez, permite supor que não há características sazonais.

Figura 5 – Comportamento temporal da variável Depósito Total da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.



Fonte: Banco Central do Brasil, organizado pelo autor.

Figura 6 – Oscilação anual da variável Depósito Total da instituição código 10045 entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.



Fonte: Banco Central do Brasil, organizado pelo autor.

Nota: Q1 – primeiro trimestre, Q2 – segundo trimestre, Q3 – terceiro trimestre e Q4 – quarto trimestre.

Após a análise gráfica, observa-se na tabela 4 que a estatística QS reporta p -valor superior a 10% para a série original (qsori) e superior a 5% para a série original ajustada

a valores extremos (qsorievadj). Neste sentido, rejeita-se a hipótese nula em favor da hipótese alternativa de ausência de sazonalidade.

Tabela 4 – Estatística QS da variável Depósito Total da instituição código 10045.

Ajuste	qs	p-valor
qsori	3.31036	0.19106
qsorievadj	5.77510	0.05571
qsrds	3.96499	0.13773
qssadj	3.31036	0.19106
qssadjevadj	5.77510	0.05571
qsirr	7.07103	0.02914
qsirrevadj	7.07103	0.02914

Fonte: Resultado da pesquisa.

Nota: qsori – série original; qsorievadj – série original corrigida por valores extremos; qsrds – resíduos do modelo ARIMA; qssadj – série temporal com ajuste sazonal; qssadjevadj – série temporal com ajuste sazonal corrigida por valores extremos; qsirr – componente irregular; qsirrevadj – componente irregular corrigido por valores extremos.

3.5.2 Raiz unitária

Além da sazonalidade é importante verificar a presença de raiz unitária nas séries temporais. Neste sentido, aplica-se os testes *Dickey-Fuller* aumentado, *Philips-Perron* e KPSS . Os testes *Dickey-Fuller* e *Philips-Perron* compartilham a hipótese nula (H_0) – a série possui raiz unitária – e hipótese alternativa (H_A) – a série é estacionária em nível –. O teste KPSS, por sua vez, é o oposto, as hipóteses são: H_0 : A série não possui raiz unitária e, H_A : A série é não é estacionária em nível.

Tabela 5 – Testes de Raiz Unitária das variáveis RIF e DT e w da instituição código 10045.

Teste	p-valor RIF	p-valor DT	p-valor w
<i>Dickey-Fuller</i> aumentado	0.8721	0.5001	0.4512
<i>Philips-Perron</i>	0.01	0.8199	0.01
KPSS	0.01	0.01	0.01

Fonte: Banco Central do Brasil.

Nota: RIF - Receita da Intermediação Financeira, DT - Depósitos Totais e w – Insumos Exógenos.

Na tabela 5 verifica-se que, para as variáveis RIF, DT e w os testes *Dickey-Fuller* aumentado e KPSS apontam a probabilidade de a série possuir raiz unitária. Já no teste *Philips-Perron*, a hipótese nula é rejeitada ao nível de 5% de significância para as variáveis

RIF e w . O mesmo não é verificado na variável DT, que indica probabilidade de a série não ser estacionária em nível no teste *Philips-Perron*.

Ademais, como se pode observar no apêndice C no diagnóstico do teste *Dickey-Fuller* aumentado, H_0 é rejeitada ao nível de significância de 1, 5 e 10% nas regressões com tendência, com *drift* e sem tendência com *drift* das três variáveis de exemplo – Receita da Intermediação Financeira (RIF), Depósito Total (DT) e Insumos Exógenos (w). Uma vez que os resultados reportados na tabela 5 e no apêndice C são uma pequena amostra do comportamento majoritário observado nas séries temporais deste estudo, conclui-se que o painel não é estacionário. Sendo assim, a regressão de dados em painel é estimada aplicando a primeira diferença em todas as variáveis.

3.5.3 Normalização logarítmica

O último tratamento realizado é a transformação logarítmica dos dados, que se faz necessário para suavizar o comportamento dos dados. Wolf (2020b) cita a proposição de Claessens e Laeven (2004) para contornar os distúrbios que a normalização logarítmica causa na variável ROA, que assume valores muito pequenos ou, em alguns casos, negativo. Assim, o *log* da variável dependente ROA é calculado de $1+ROA$. O mesmo é aplicado a variável CRD, que também assume valores muito baixo e em poucos casos negativos.

Logo, o modelo empírico assume a forma:

$$\ln RIF_{it} = \alpha + \gamma_1 \ln DT_{it} + \beta \ln w_{it} + \gamma_2 \ln CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (18)$$

$$\ln RT_{it} = \alpha + \gamma_1 \ln DT_{it} + \beta \ln w_{it} + \gamma_2 \ln CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (19)$$

$$\ln ROA_{it} = \alpha + \gamma_1 \ln DT_{it} + \beta \ln w_{it} + \gamma_2 \ln CRD_{it} + \gamma_3 TB_{it} + \epsilon_{it} \quad (20)$$

3.5.4 Dados faltantes

Da dessazonalização e a normalização logarítmica, gerou-se algumas observações do tipo: números faltantes, não numéricos e infinitos. Logo, faz-se necessário novos tratamentos para se obter um banco de dados computável. Neste ponto, considera-se importante manter o maior número de observações e, também, a integridade da série temporal de cada instituição financeira. Sendo assim, os valores ausentes e não numéricos foram substituídos pela mediana. Porém, quando a ocorrência deste valores é observada

no início ou final da série temporal, opta-se por eliminar as primeiras ou últimas linhas de observações. É importante lembrar que, as variáveis Depósito Total (DTT), Saldo das Operações de Crédito (SOC) e Ativo Total Ajustado (ATA) do modelo $IHHn$, não recebem tratamento logarítmico e sazonais. O seu tratamento se faz pela exclusão das observações com números ausentes. Já para as variáveis que são usadas na estimativa do H de Panzar-Rosse e para as curvas das *proxies* de custo médio e preço do crédito, os tratamentos seguem caso a caso.³⁴

Em comparação com a tabela 2 observa-se que com os tratamentos o banco de dados sofre uma redução de 5 observações. Contudo, verifica-se na tabela 6 que as variáveis CRD, PC e CM não apresentam mais média e máxima tendendo ao infinito. Ou seja, foi corrigido o problema de valores explosivos. Logo, é possível com os dados obter estimações plausíveis da estatística H de Panzar-Rosse e traçar as curvas *proxies* do custo médio e preço do crédito.

³⁴ Ver: Apêndice D.

Tabela 6 – Estatística descritiva das variáveis dessazonalizadas, com transformação logarítmica e os números ausentes tratados.

Variável	N	Média	Desvio Padrão	Mínima	1º Quartil	3º Quartil	Máxima
RIF	6,741	12.498	2.217	5.686	10.943	13.873	18.396
RT	6,741	12.736	2.198	6.906	11.203	14.056	18.558
ROA	6,741	0.021	0.132	-3.937	0.002	0.019	1.492
DT	6,741	13.766	3.030	-2.132	12.270	15.455	20.373
w	6,741	12.219	2.147	6.992	10.692	13.513	18.010
CRD	6,741	0.892	0.583	0.00002	0.582	1.111	10.698
TB	6,741	0.008	0.087	0	0	0	1
DTT	6,661	23,829,117.000	90,078,694.000	0.000	229,246.600	5,292,027.000	760,018,182.000
SOC	6,741	23,021,288.000	89,182,873.000	13.898	254,508.100	6,514,999.000	812,113,751.000
ATA	6,741	63,749,185.000	222,361,741.000	21,642.280	847,362.000	19,263,459.000	1,723,632,154.000
PC	6,741	0.418	0.488	-0.398	0.165	0.492	9.349
CM	6,741	0.475	0.538	-0.675	0.194	0.553	10.268

Fonte: Resultados da pesquisa.

Nota: RIF – Receita da Intermediação Financeira; RT – Receita Total; ROA – *Return of Asset*; DT – Depósito Total; w – Insumos Exógenos (Despesas de Pessoal + Despesas Administrativas + Despesas de Captação); CRD – Risco da Intermediação Financeira; TB – Tipo Bancário; DTT – Depósitos Totais (DT – Depósitos da Intermediação Financeira + Letras de Crédito Imobiliário + Letras de Crédito do Agronegócio + Letras Financeiras); SOC – Saldo das Operações de Crédito; ATA – Ativo Total Ajustado; PC – Preço do Crédito e CM – Custo Médio.

3.5.5 Matriz de correlação

Na tabela 7 é apresentada a Matriz de correlação das variáveis – dessazonalizadas e com transformação logarítmica – utilizadas para a estimação da estatística H de Panzar-Rosse e para a elaboração das curvas de análise do Custo Médio e Preço do Crédito. Observa-se que a variável RIF reporta valores do coeficiente de *Pearson* (r)³⁵ muito próximos aos da variável RT. Contudo, a variável ROA apresenta valores $r = 0$ e muito próximo de zero em todas as suas interações. Também são interessantes de se ressaltar: a correlação moderada e positiva entre o número de agências bancárias físicas (NAG) e as RIF (0,57) e RT (0,58); a baixa correlação da variável TB com todas as demais variáveis; a alta correlação positiva entre o PC e o CM (0,93); e a correlação positiva e não baixa (0,69) entre as variáveis independentes no modelo da estatística H de Panzar-Rosse w e DT.

Tabela 7 – Matriz de correlação das variáveis tratadas, dessazonalizadas e loglinearizadas.

	RIF	RT	ROA	DT	w	CRD	TB	PC	CM	NAG
RIF	1	0,99	0	0,67	0,96	-0,03	-0,02	-0,06	-0,13	0,57
RT	0,99	1	0	0,68	0,97	-0,05	-0,02	-0,06	-0,09	0,58
ROA	0	0	1	0	-0,03	0,03	0	0,03	-0,03	-0,03
DT	0,67	0,68	0	1	0,69	-0,21	0	-0,16	-0,18	0,44
w	0,96	0,97	0	0,69	1	-0,06	0	-0,07	-0,08	0,59
CRD	-0,03	-0,05	0,03	-0,21	-0,06	1	-0,05	-0,06	-0,09	-0,03
TB	-0,02	-0,02	0	0,02	0	-0,05	1	-0,04	-0,02	-0,03
PC	-0,06	-0,06	0,03	-0,16	-0,07	-0,06	-0,04	1	0,93	0,10
CM	-0,13	-0,09	-0,03	-0,18	-0,08	-0,09	-0,02	0,93	1	-0,09
NAG	0,57	0,58	-0,03	0,44	0,59	-0,03	-0,03	-0,10	-0,09	1

Fonte: Resultados da pesquisa.

Nota: RIF – Receita da Intermediação Financeira; RT – Receita Total; ROA – *Return of Asset*; DT – Depósito Total; w – Insumos Exógenos (Despesas de Pessoal + Despesas Administrativas + Despesas de Captação); CRD – Risco da Intermediação Financeira; TB – Tipo Bancário; PC – Preço do Crédito e CM – Custo Médio.

³⁵ O coeficiente de *Pearson* (r) é definido por:

$$r = \frac{1}{1-n} \sum_1^n \frac{(x_i - \bar{x})}{sd(x)} \frac{(y_i - \bar{y})}{sd(y)}$$

restrito ao intervalo entre -1 e 1 .

4 RESULTADOS

Nesta seção são exibidos a análise exploratória de dados espaciais, os resultados da estimativa do IHH_n , os gráficos das *proxies* do custo médio, do preço do crédito e as estimativas das estatísticas H de Panzar-Rosse.

4.1 ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS

Nesta seção apresenta-se a motivação do trabalho com o auxílio de instrumentos de análise exploratória dos dados espaciais (AEDE). A AEDE é um conjunto de técnicas que esboçam a distribuição dos dados no espaço geográfico. Os dados utilizados são da Estatística Bancária³⁶ de dezembro de 2019 do BACEN, a densidade populacional estimada pelo Instituto Brasileiro de Geografia e Estatística (IBGE)³⁷ e os *shapfiles* da malha municipal de 2020³⁸, das localidades e limites de 2017³⁹.

4.1.1 Análise municipal

Na média, considerando os 5.570 municípios do Brasil no ano de 2019, existe uma agência bancária para cada grupo de 10.551 pessoas. Se forem consideradas somente as cidades que possuem ao menos uma unidade de atendimento presencial, este valor cai para quase 9.694 moradores por agência. Na tabela 8 estão as 10 cidades de maior densidade populacional, ordenadas da maior para a menor, em que se observa a estimativa populacional, o número de agências bancárias e a estimativa do número de moradores por agência. Por meio desta, verifica-se que somente Salvador (BA) está acima das duas médias supracitadas e Porto Alegre (RS) é a cidade com a menor proporção de habitantes por agências, 4.429.

³⁶ Relatório por municípios. Disponível em: <https://www4.bcb.gov.br/fis/cosif/estban.asp?frame=1>

³⁷ Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/populacao>

³⁸ Disponível em: <https://portaldemapas.ibge.gov.br/portal.php>

³⁹ Disponíveis em: <https://www.ibge.gov.br/geociencias/downloads-geociencias.html>

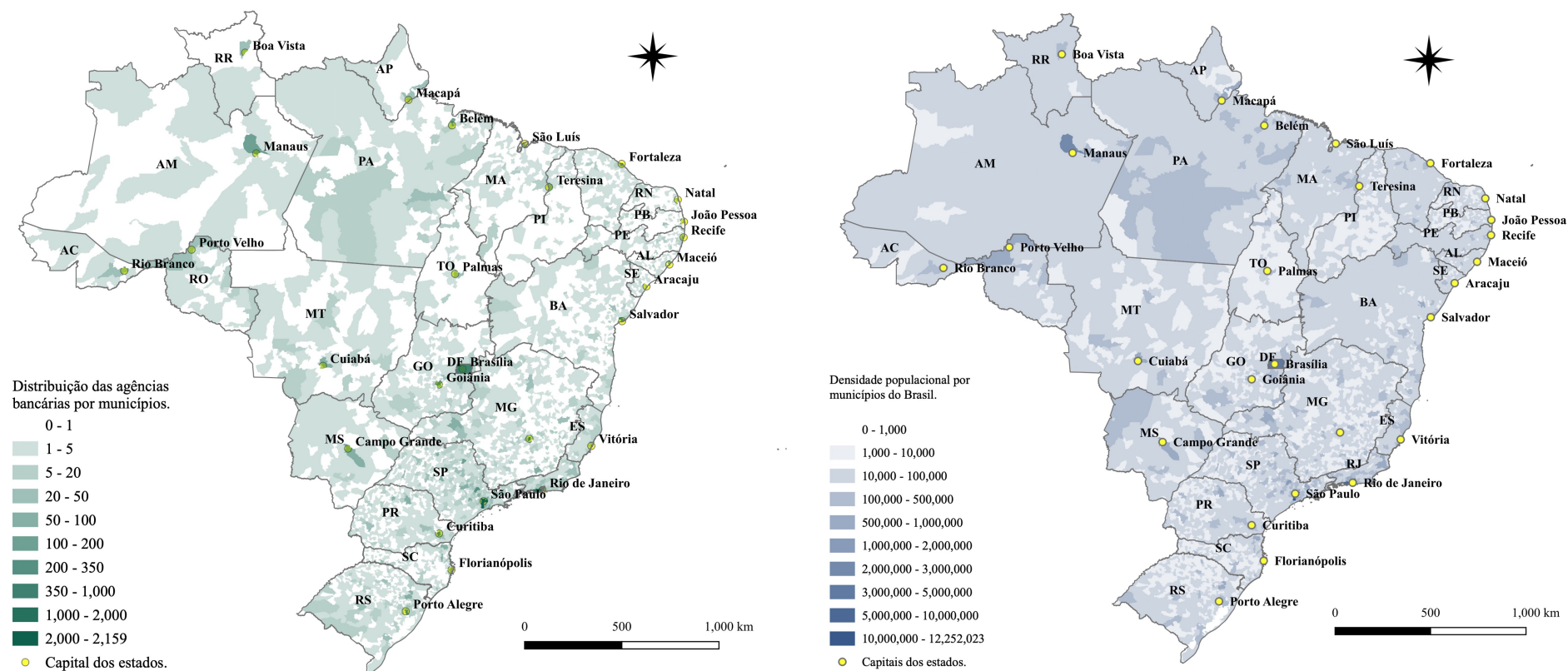
Tabela 8 – Análise dos 10 municípios brasileiros com a maior densidade populacional em 2019.

Municípios	Ufs	Est. Populacional	Nº Agências	Habitantes por agências
São Paulo	SP	12.252.023	2.159	5.674,86
Rio de Janeiro	RJ	6.718.903	1.027	6.542,26
Brasília	DF	3.015.268	411	7.336,42
Belo Horizonte	MG	2.512.070	386	6.507,95
Curitiba	PR	1.933.105	346	5.587,01
Porto Alegre	RS	1.483.771	335	4.429,17
Salvador	BA	2.872.347	256	11.220,11
Campinas	SP	1.204.073	223	5.399,43
Goiânia	GO	1.516.113	203	7.468,54
Recife	PE	1.645.727	192	8.571,49

Fonte: Banco Central do Brasil e Instituto Brasileiro de Geografia e Estatística (IBGE), organizado pelo autor.

Na figura 7 observa-se a disparidade na distribuição dos serviços bancários ofertados no Brasil em dezembro de 2019. São 2.331 cidades pintadas na cor branca, representando os municípios onde não existem agências bancárias. São aproximadamente 17.052.647 pessoas sem acesso físico a serviços financeiros básicos em suas cidades. Adicionando a este grupo as pessoas cuja cidade é servida por uma agência – 1.077 municípios –, o número cresce para 30.273.324 indivíduos. Isto significa que, cerca de 14,41% da população brasileira reside em cidades com apenas uma ou nenhuma agência bancária física.

Figura 7 – Distribuição espacial das agências bancárias e da densidade demográfica dos municípios brasileiros em dezembro de 2019.



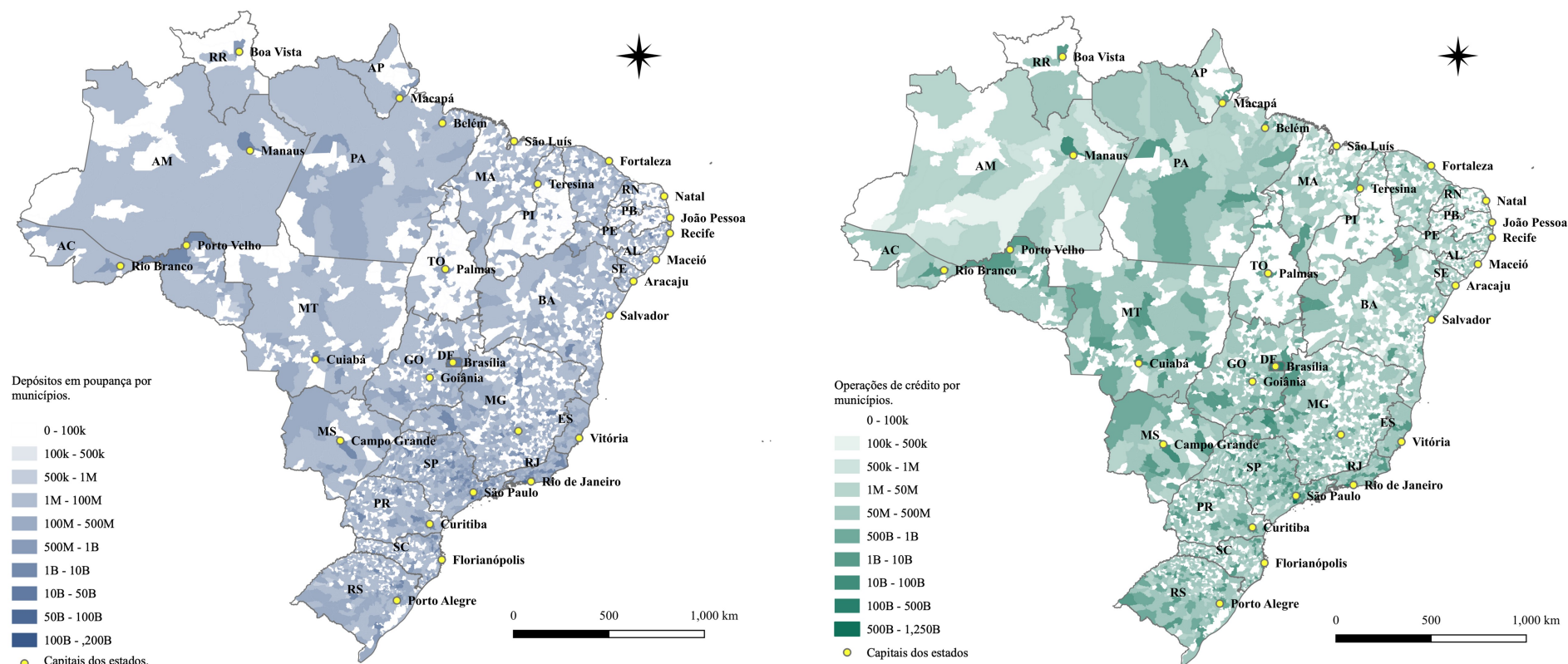
Fonte: Banco Central do Brasil e Instituto Brasileiro de Geografia e Estatística (IBGE), organizado pelo autor..

Quando analisamos as operações de crédito por municípios, fica mais forte o contraste entre as cidades brasileiras. São Paulo (SP) é a cidade com o maior volume de operações de crédito, pouco mais que um trilhão e duzentos e dez bilhões de reais⁴⁰. A segunda cidade com o maior volume de operações de crédito é Osasco (SP), e o volume é 3,4 vezes menor que o da cidade de São Paulo. Juntas as duas cidades possuem um volume maior que 1,5 trilhões de reais. A soma das operações de crédito do resto do Brasil não chega a 1,7 trilhões de reais. Ou seja, sozinha a cidade de São Paulo (SP) realiza 37,12% de todas as operações de crédito efetuadas no país. Quando somada a Osasco (SP), esta porcentagem sobe para cerca de 48%. Em ordem, as 10 cidades com os maiores volumes de operações de crédito e a porcentagem são: São Paulo SP 37,11%, Osasco SP 10,91%, Brasília (DF) 4,77%, Rio de Janeiro (RJ) 3,70%, Porto Alegre (RS) 2,34%, Belo Horizonte (MG) 2,26%, Curitiba (PR) 1,55%, Goiânia (GO) 1,03%, Fortaleza (CE) 0,84% e Barueri (SP) 0,84%. Desta forma, 65,36% do volume de operações de crédito realizados no período, está concentrado em 15,74% da população brasileira estimada pelo IBGE. (Figura 8).

A distribuição muda quando se analisa os volumes dos depósitos em poupança. A cidade de São Paulo (SP) continua sendo a que possui o maior volume porém, agora a diferença para a segunda cidade – Rio de Janeiro (RJ) – é de 1,63 vezes. Quando comparada com a soma dos depósitos em poupança do resto do país, o volume da capital paulista representa para o todo 12,44%. Em ordem, as 10 cidades com o maior volume de depósitos em poupança com sua porcentagem na participação nacional são: São Paulo SP 12,44%, Rio de Janeiro (RJ) 7,64%, Porto Alegre (RS) 3,46%, Brasília (DF) 2,68%, Belo Horizonte (MG) 2,52%, Curitiba (PR) 1,94%, Salvador (BH) 1,59%, Fortaleza (CE) 1,29%, Recife (PE) 1,28% e Goiânia (GO) 1,07%. Somadas, as 10 cidades representam 35,91% de todo o volume de depósitos em poupança dos 5570 municípios. Ou seja, pouco mais que 17% da estimativa populacional do Brasil concentra quase 36% do total de depósitos para o período. (Figura 8).

⁴⁰ R\$ 1.210.359.769.211,00

Figura 8 – Distribuição espacial das operações de crédito e dos depósitos em poupança dos municípios brasileiros em dezembro de 2019.



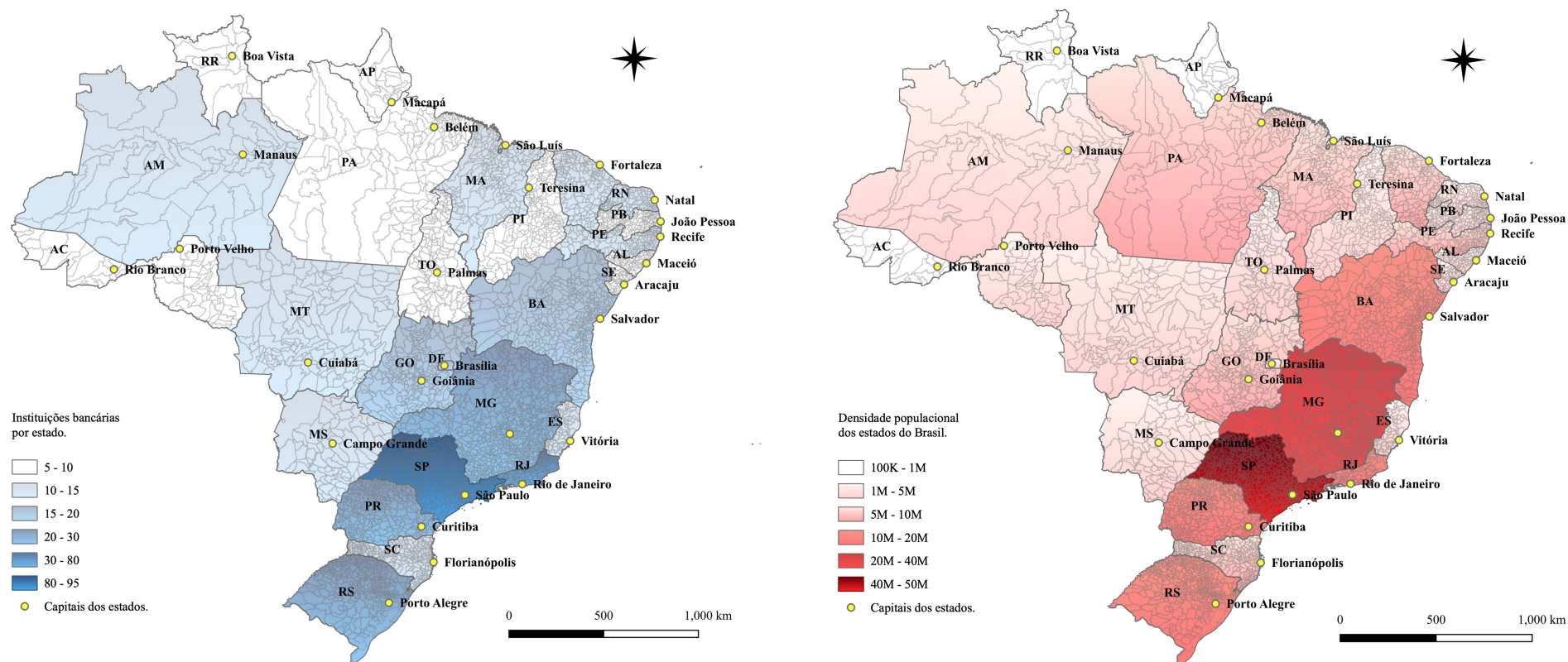
Fonte: Banco Central do Brasil e Instituto Brasileiro de Geografia e Estatística (IBGE), organizado pelo autor..

4.1.2 Análise estadual

O Centro-oeste é a região nacional com a menor densidade populacional, com aproximadamente 8% do total de habitantes do Brasil. Ao mesmo tempo, a região concentra 8% do total de agências bancárias – 1659 –, o que representa na média uma unidade de atendimento físico para cada grupo de 9.823 indivíduos. Em contrapartida, o Nordeste, com uma estimativa populacional em 2019 de mais de 57 milhões de habitantes, concentra mais que 16% das agências bancárias nacionais, tendo em média uma agência bancária para cada grupo de 17.571 moradores. A região sul, que abriga cerca de 14% da população brasileira, possui cerca de 19% das agências bancárias e a melhor média agência por habitantes com cerca de 8.071 moradores para cada unidade bancária. Já nas regiões Sudeste, estima-se, em média, uma agência bancária para cada grupo de 8.665 e 9.823 indivíduos, respectivamente.

Quando se observa o número de instituições bancárias que atuam em cada estado, figura 9 e tabela 9, a diferença fica evidente. Enquanto os estados de São Paulo, Rio de Janeiro e Minas Gerais, possuem respectivamente 95, 36 e 29 instituições bancárias atuantes, Acre, Rondônia e Roraima possuem apenas 6. Em dezembro de 2019, 9 unidades federativas contavam com menos de 10 instituições, representados pela cor branca. Por sua vez, apenas o estado de São Paulo contava com mais de 40 firmas, sendo representado pelo azul mais escuro. Contudo, quando se compara a densidade populacional dos estados com as presenças de instituições bancárias, Roraima é o destaque positivo com, em média, uma firma para cada grupo de 101 mil habitantes. O Pará ressalta negativamente com a relação de uma instituição para cerca de 860 mil habitantes. Na média, a distribuição da razão da estimativa populacional pelo número de instituições atuantes em cada estado é de um banco para cada grupo de aproximadamente 411 mil pessoas. Contudo, a razão entre a estimativa populacional nacional – 210.147.175 – e o número de instituições do relatório EstBan por municípios de dezembro de 2019 – 113 – é de uma empresa para aproximadamente 1.860.000 habitantes.

Figura 9 – Distribuição espacial do número de instituições bancárias e, da estimativa populacional brasileiros em dezembro de 2019.



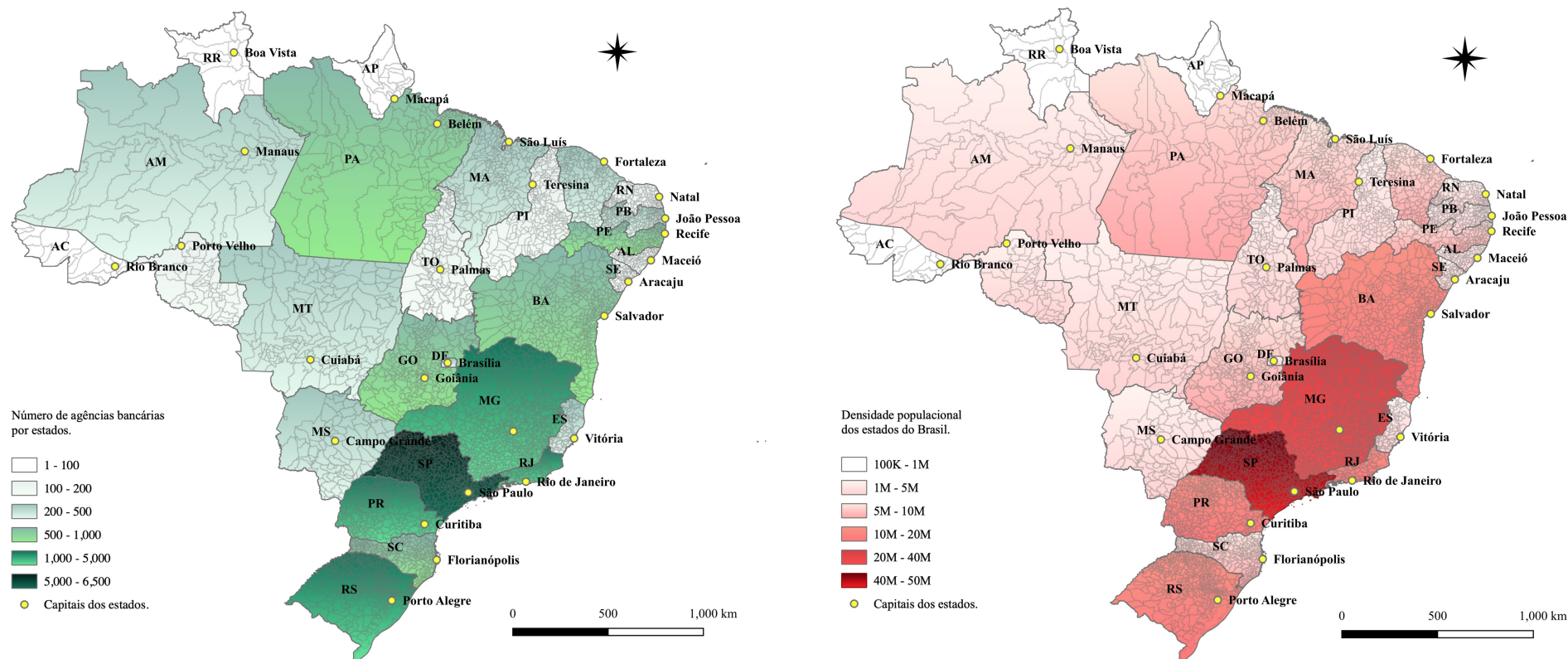
Fonte: Banco Central do Brasil e Instituto Brasileiro de Geografia e Estatística (IBGE), organizado pelo autor.

No mapa, figura 10, fica claro que a quantidade de agências e a densidade populacional tende a se concentrar à medida que segue do Norte para o Sudeste. Também, verifica-se que dos estados do Sul e Sudeste, somente Santa Catarina e o Espírito Santo se destacam por estar fora dos “segundos e terceiros” tons mais escuros, tanto no mapa correspondente ao número de agências bancárias por estado, quanto no mapa da estimativa da densidade populacional. Ao mesmo tempo, destacam-se os 3 estados do Norte, Roraima, Acre e Amapá, coloridos de branco nos dois mapas, por possuírem as menores estimativas populacionais e o menor número de agências bancárias. São Paulo se destaca por ser o estado que possui a maior densidade populacional e, também, o maior número de agências, coloridos nos tons mais escuros. Na comparação com o segundo estado – Minas Gerais –, SP possui a densidade populacional cerca de 2,2 vezes maior e quantidade de agências 3,1 vezes maior.

De acordo com a estatística bancária mensal por municípios, de dezembro de 2019, são 19.917 agências bancárias no Brasil⁴¹. Porém, a distribuição não é nada harmônica. Enquanto no estado de São Paulo há 6.048 unidades físicas de atendimento, em Roraima são apenas 30. Em termos populacionais, é uma agência para, em média, 20.192 habitantes no estado de Roraima enquanto em São Paulo a estimativa é de em média 7.592. Contudo, seguindo esta métrica, o destaque positivo é do estado do Rio Grande do Sul que tem em média 7.097 indivíduos para cada agência bancária. Negativamente, o estado do Maranhão se sobressai com uma unidade física de atendimento para cada 22.041 habitantes. Na tabela 9, um outro dado deve ser enfatizado: a proporção de operações de crédito em relação aos depósitos em poupança. Neste contexto, o estado do Rio de Janeiro tem o destaque por ser o estado com a maior proporção de direcionamento dos recursos captados nos depósitos de poupança para as operações de crédito. Negativamente, São Paulo, Distrito Federal e o estado do Mato Grosso possuem a menor taxa, menos de 15,5% da poupança é destinada para as operações de crédito. A média nacional dos estados é 38,06%, ou seja, por esta métrica, tende-se a verificar o resultado propostos por Belaisch (2003) e Lucinda (2010). Os bancos, principalmente os grandes bancos, tendem a ignorar o mercado de crédito de varejo e emprestar para o governo, comprando títulos públicos, e, com isto, reduzir sua exposição ao risco.

⁴¹ No relatório da estatística bancária existem duas colunas para a quantidade de agências por instituição por cidades. Seguem os somatórios e as definições para as duas diferentes colunas.
 19917 – AGEN_PROCESSADAS: apresentado somente no arquivo por município, é a quantidade de agências processadas da Instituição Financeira no município.
 19926 – AGEN_ESPERADAS: apresentado somente no arquivo por município, é a quantidade de agências esperadas da Instituição Financeira no município.

Figura 10 – Distribuição espacial do número de instituições bancárias e, da estimativa populacional brasileiros em dezembro de 2019.



Fonte: Banco Central do Brasil e Instituto Brasileiro de Geografia e Estatística (IBGE), organizado pelo autor.

Tabela 9 – Análise comparativa de dados bancários com a população estimada por estados para o ano de 2019.

Unidade Federativa	Ufs	Região	População (a)	% de (a)	Nº Inst. (b)	Nº Agências (c)	% de (c)	(a) / (c)	Oper. de Crédito (d)	Depo. em Poupança (e)	(e) / (d)
Rondônia	RO	Norte	1.777.225	0,85%	6	138	0,69%	12.878,44	130.11.286.012	3.499.657.814	0,27
Acre	AC	Norte	881.935	0,42%	6	56	0,28%	15.748,84	4.135.033.011	1.283.288.416	0,31
Amazonas	AM	Norte	4.144.597	1,97%	11	206	1,03%	20.119,40	13.353.018.354	5.369.988.669	0,40
Roraima	RR	Norte	605.761	0,29%	6	30	0,15%	20.192,03	3.699.194.863	872.217.067	0,24
Pará	PA	Norte	8.602.865	4,09%	10	511	2,57%	16.835,35	28.985.461.829	11.016.826.307	0,38
Amapá	AP	Norte	845.731	0,40%	8	43	0,22%	19.668,16	3.783.818.220	935.670.709	0,25
Tocantins	TO	Norte	1.572.866	0,75%	8	114	0,57%	13.797,07	11.629.801.630	2.770.464.459	0,24
Maranhão	MA	Nordeste	7.075.181	3,37%	12	321	1,61%	22.041,06	23.069.190.624	9.547.959.108	0,41
Piauí	PI	Nordeste	3.273.277	1,56%	7	163	0,82%	20.081,45	15.137.833.490	7.100.638.521	0,47
Ceará	CE	Nordeste	9.132.078	4,35%	14	461	2,31%	19.809,28	40.205.728.094	19.263.938.869	0,48
Rio Grande do Norte	RN	Nordeste	3.506.853	1,67%	11	190	0,95%	18.457,12	18.188.934.986	7.226.724.058	0,40
Paraíba	PB	Nordeste	4.018.127	1,91%	9	225	1,13%	17.858,34	18.720.472.564	9.119.247.437	0,49
Pernambuco	PE	Nordeste	9.557.071	4,55%	15	542	2,72%	17.632,97	40.066.511.118	23.106.446.376	0,58
Alagoas	AL	Nordeste	3.337.357	1,59%	9	176	0,88%	18.962,26	13.671.507.672	6.456.962.266	0,47
Sergipe	SE	Nordeste	2.298.696	1,09%	9	193	0,97%	11.910,34	13.968.644.489	6.737.676.330	0,48
Bahia	BA	Nordeste	14.873.064	7,08%	19	977	4,91%	15.223,20	61.972.969.844	37.262.766.367	0,60
Minas Gerais	MG	Sudeste	21.168.791	10,07%	29	1926	9,67%	10.991,06	191.764.810.396	88.797.492.829	0,46
Espírito Santo	ES	Sudeste	4.018.650	1,91%	11	411	2,06%	9.777,74	28.126.211.715	14.988.250.922	0,53
Rio de Janeiro	RJ	Sudeste	17.264.943	8,22%	36	1813	9,10%	9.522,86	162.767.945.255	104.815.753.895	0,64
São Paulo	SP	Sudeste	45.919.049	21,85%	95	6048	30,37%	7.592,44	1.860.270.278.136	267.953.865.995	0,14
Paraná	PR	Sul	11.433.957	5,44%	27	1344	6,75%	8.507,41	135.647.719.610	49.217.007.966	0,36
Santa Catarina	SC	Sul	7.164.788	3,41%	15	767	3,85%	9.341,31	75.419.405.236	34.517.430.803	0,46
Rio Grande do Sul	RS	Sul	11.377.239	5,41%	23	1603	8,05%	7.097,47	168.541.778.257	73.964.129.350	0,44
Mato Grosso do Sul	MS	Centro-oeste	2.778.986	1,32%	12	258	1,30%	10.771,26	31.486.697.085	6.479.980.667	0,21
Mato Grosso	MT	Centro-oeste	3.484.466	1,66%	14	307	1,54%	11.350,05	47.938.733.568	7.344.997.766	0,15
Goiás	GO	Centro-oeste	7.018.354	3,34%	18	683	3,43%	10.275,77	79.716.179.426	21.187.467.036	0,27
Distrito Federal	DF	Centro-oeste	3.015.268	1,43%	17	411	2,06%	7.336,42	155.614.245.043	22.590.662.431	0,15
Totais (T) e médias (M)		Brasil	210.147.175 (T)	1 (T)	112 (T)	19917 (T)	1 (T)	14.214,04 (M)	3.260.893.410.527 (T)	843.427.512.433 (T)	38,06% (M)

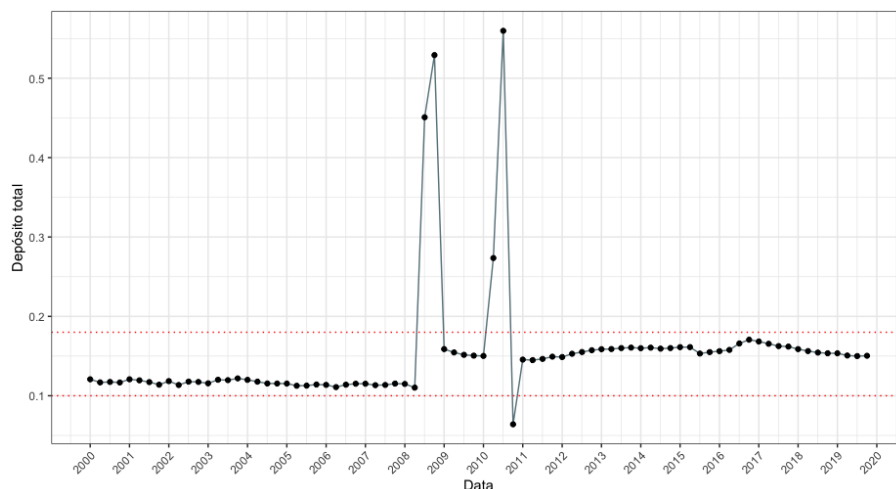
Fonte: Banco Central do Brasil e Instituto Brasileiro de Geografia e Estatística (IBGE), organizado pelo autor.

4.2 ÍNDICE DE *HERFINDAHL-HIRSCHMAN* NORMALIZADO

Os resultados obtidos na estimação do índice de *Herfindahl-Hirschman* normalizado ($IHHn$) para os Depósitos Totais (DTT), os Saldos das Operações de Crédito (SOC) e os Ativos Totais Ajustados (ATA), estão dispostos nos gráficos a seguir. No apêndice E, é reportado a tabela utilizada para produzir estes gráficos. Além disso, na mesma tabela também são reportados o número de observações – que representa o número de instituições bancárias da amostra em cada trimestre – e o número de agências bancárias totais por período e os valores médios por período dos custos médios bancários e dos preços do crédito.

Os produtos das estimações têm resultados próximos aos reportados pelo BACEN nos Relatórios da Economia Bancária, sendo as diferenças raramente superiores a 0.78%. Tal resultado reforça a qualidade da amostra deste ensaio. Outro ponto relevante é que os resultados também se mostraram próximos aos reportados por Wolf (2020a). Contudo, é importante destacar que a variável Depósito Total (DTT) reporta quatro valores discrepantes no terceiro e quarto trimestre de 2008 e no segundo e terceiro trimestre de 2010. Na análise dos resultados, seguindo os critérios do BACEN, verifica-se que, fora os quatro casos supracitados nos depósitos totais, a indústria bancária brasileira oscila entre níveis de concentração baixos e moderados até o quarto trimestre de 2010. A partir do primeiro trimestre de 2011, até o quarto trimestre de 2019, os resultados das 3 variáveis oscilam dentro das fronteiras dos níveis de concentração moderado.

Figura 11 – $IHHn$ dos Depósitos Totais entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.



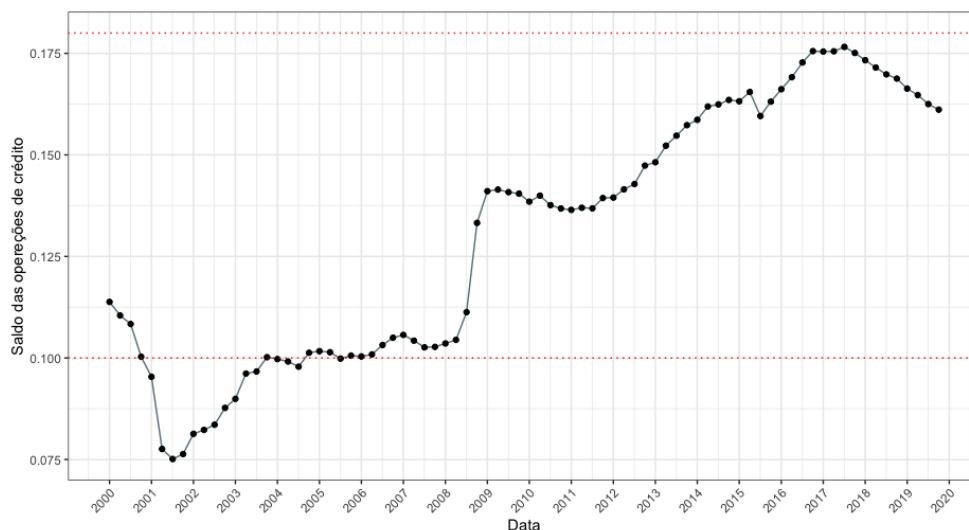
Fonte: Resultados da pesquisa.

Nota: A linha pontilhada vermelha representa os intervalos entre os níveis de concentração. Segundo o BANCO CENTRAL DO BRASIL (2019b), os níveis de concentração são considerados baixos se entre 0 e 0,1000; moderados se maior que 0,1000 e menor que 0,1800; e elevados se maior que 0,1800 até 1.

A figura 11 reporta a comportamento trimestral do nível de concentração da variável Depósitos Totais (DTT). Observa-se que em 2008 e 2010 o índice ultrapassa as linhas pontilhadas vermelhas que representam os limites superior e inferior de concentração estipulados pelo Banco Central do Brasil. No terceiro trimestre de 2008, o $IHHn$ reporta valor de 0,4508 e 0,5291 no acumulado do próximo trimestre. Não foi encontrado explicação na literatura pesquisada para o resultado obtido. Contudo, observa-se uma quantidade expressiva de dados faltantes nos relatórios bancários destas datas. Não aparenta ser imprudente creditar à crise dos *subprimes* o deslocamento ou queda acentuada nas aplicações em poupança realizada pelos agentes econômicos. Uma vez que o sistema bancário e financeiro teve sua credibilidade abalada após a repentina quebra de grandes bancos, verifica-se ser plausível que os agentes econômicos tendam a alocar seus recursos em outros setores da economia menos voláteis naquele período.

Para as oscilações, a maior – 0,2735 no segundo trimestre e 0,5597 no terceiro trimestre – e a menor – 0,0640 quarto trimestre – apresentadas em 2010, também não foi encontrada uma explicação concreta para a piora nos níveis de concentração. Porém, observa-se quantidades expressivas de dados ausentes nos maiores bancos privados do país nestes períodos.

Figura 12 – $IHHn$ dos Saldos das Operações de Crédito entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.

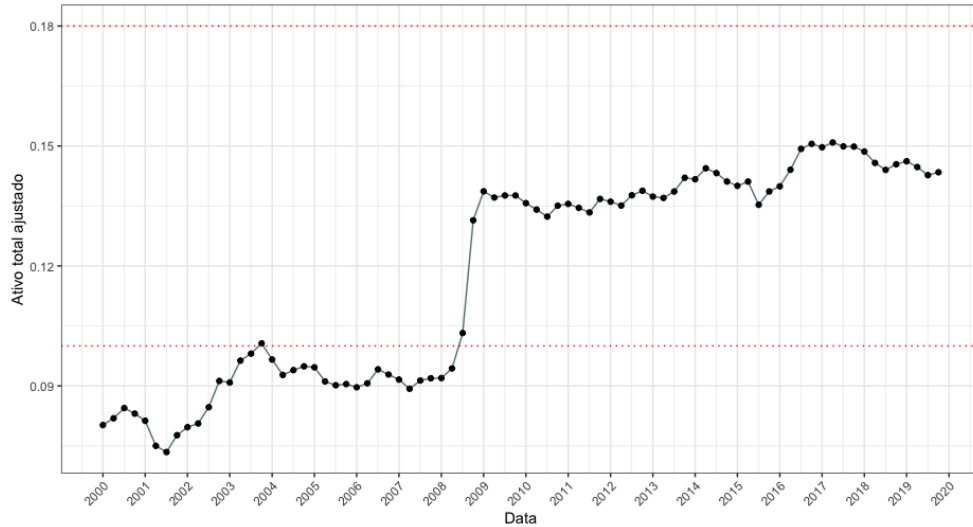


Fonte: Resultados da pesquisa.

Nota: A linha pontilhada vermelha representa os intervalos entre os níveis de concentração. Segundo o BANCO CENTRAL DO BRASIL (2019b), os níveis de concentração são considerados baixos se entre 0 e 0,1000; moderados se maior que 0,1000 e menor que 0,1800; e elevados se maior que 0,1800 até 1.

O gráfico da figura 12 mostra o comportamento do Índice de *Herfindahl-Hirschman* normalizado para a variável Saldo das Operações de Crédito (SOC). Observa-se que entre 2000 e 2003 os níveis de concentração estiveram abaixo de 0,100. Porém, voltam para a área de níveis moderados em uma ascendente que vai até os anos de 2016 e 2017, chegando próximo a barreira dos 0,180 que limita o índice entre moderado e elevado nível de concentração. Assim como para a variável Depósito Total, a variável SOC aparenta ter uma tendência de redução do $IHHn$ a partir dos anos de 2016 e 2017, quando passam a vigorar as medidas de regulamentação que possibilitam e incentivam a indústria bancária a se digitalizar.

Figura 13 – $IHHn$ dos Ativos Totais Ajustados entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.



Fonte: Resultados da pesquisa.

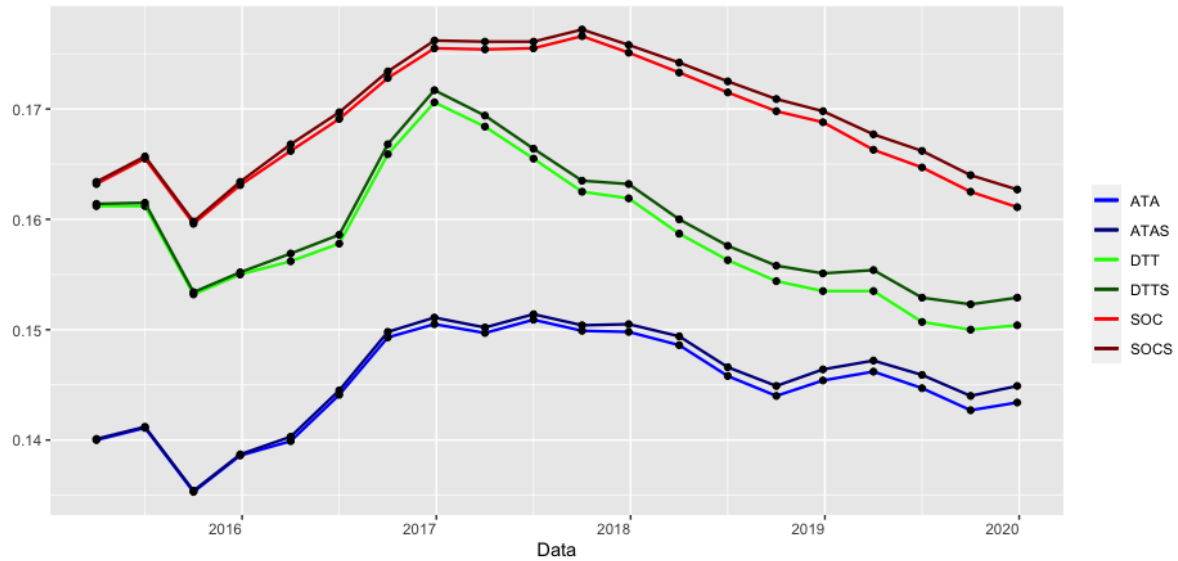
Nota: A linha pontilhada vermelha representa os intervalos entre os níveis de concentração. Segundo o BANCO CENTRAL DO BRASIL (2019b), os níveis de concentração são considerados baixos se entre 0 e 0,1000; moderados se maior que 0,1000 e menor que 0,1800; e elevados se maior que 0,1800 até 1.

Na figura 13 – $IHHn$ do Ativo Total Ajustado (ATA) – verifica-se que em 2008 houve um salto a maior tirando a indústria de níveis baixos de concentração para níveis moderados de concentração. Credita-se este movimento à grande quantidade de fusões e aquisições observadas no período. Como exemplo, pode-se citar a fusão do Itaú com o Unibanco e aquisição do ABN Amro pelo Santander.

4.2.1 $IHHn$ com e sem os bancos digitais

A análise a seguir compara a estimativa do Índice *Herfindahl-Hirschman* normalizado, no período entre o primeiro trimestre de 2015 até o quarto trimestre de 2019, calculado das variáveis Depósito Total (DTT), Saldo das Operações de Crédito (SOC) e Ativo Total Ajustado (ATA) em duas amostras, com e sem os bancos digitais. A amostra das variáveis sem os bancos digitais recebe a letra S na sigla – DTTS, SOCS e ATAS – e são representadas pelas cores mais escuras. Já a amostra com os bancos digitais permanece com a sigla já utilizada neste ensaio e são representadas pelas cores claras – DTT (verde claro), SOC (vermelho claro) e ATA (azul claro).

Figura 14 – $IHHn$ das variáveis Depósito Total, Saldo das Operações de Crédito e Ativo Total Ajustado do primeiro trimestre de 2015 ao quarto trimestre de 2019.



Fonte: Resultados da pesquisa.

Nota: DTT – Depósito Total, DTTS – Depósito Total sem os bancos digitais, SOC – Saldo das Operações de Crédito, SOCS – SOC – Saldo das Operações de Crédito sem os bancos digitais, ATA – Ativo Total Ajustado e ATAS – Ativo Total Ajustado sem os bancos digitais.

A figura 14 mostra que, apesar de a diferença ser pequena dos níveis de concentração da amostra sem os bancos digitais para a estimativa com os bancos digitais, ela é crescente. Observa-se que na data inicial da comparação – primeiro trimestre de 2015 – os pares de linhas das seis variáveis, iniciam a trajetória sobrepostas. A sobreposição começa a se desfazer no primeiro trimestre de 2016, a partir deste ponto, averigua-se um afastamento contínuo entre as curvas e seus pares. Porém, sem que haja alteração na trajetória das curvas. Verifica-se, também, que as três variáveis com os bancos digitais, apresentam níveis de concentração menores que as variáveis na qual os bancos digitais não estão presentes na indústria.

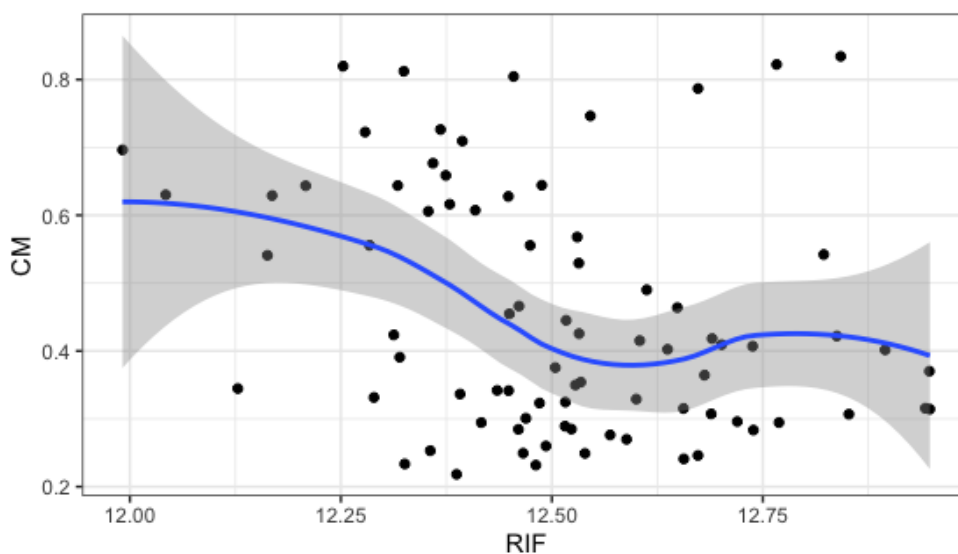
A variável Depósito Total (DTT) se destaca, pois, comparando com as variáveis Saldo das Operações de Crédito (SOC) e Ativo Total Ajustado (ATA), é a que reporta maior diferença entre os níveis de concentração, no quarto trimestre de 2019, nas especificações sem os bancos digitais (DTTS, SOCS e ATAS) e com os bancos digitais (DTT, SOC, ATA) – $IHHn^{DTTS} = 0.1529$, $IHHn^{DTT} = 0.1504$, $IHHn^{SOCS} = 0.1627$, $IHHn^{SOC} = 0.1611$, $IHHn^{ATAS} = 0.1449$ e $IHHn^{ATA} = 0.1434$. Isto significa dizer que, a estratégia empírica adotada é eficaz em indicar que a introdução dos bancos digitais na indústria bancária nacional contribui para a redução dos níveis de concentração do setor.

4.3 CURVAS DE CUSTO MÉDIO E PREÇO DO CRÉDITO

As curvas analíticas apresentadas nesta subseção são usadas para validar o uso da estatística H , haja vista as críticas feitas por Spierdijk e Shaffer (2015) e Bikker, Shaffers e Spierdijk (2012).

A análise da curva do gráfico (desenhado a partir da média por trimestre do Custo Médio ($\overline{CM}$) – no eixo y – em relação a média por trimestre da Receita da Intermediação Financeira ($\overline{RIF}$)⁴² – no eixo x) mostra o comportamento descendente da $\overline{CM}$ em relação ao aumento da $\overline{RIF}$. A inclinação reflete a suposição de que exista economia de escala na indústria bancária brasileira, de modo que, quanto maior for produção de crédito (medido pela RIF), menores tendem a serem os custos por unidade. Krugman e Obstfeld (2001).

Figura 15 – Gráfico de dispersão das médias trimestrais do Custo Médio e em relação a Receita da Intermediação Financeira..



Fonte: Resultados da pesquisa.

Nota: RIF – Receita da Intermediação Financeira e CM – Custo Médio. Valores em log. A linha azul é calculada via Mínimos Quadrados Ordinários, o sombreado cinza que envolve a linha azul representa o intervalo de confiança de 95%.

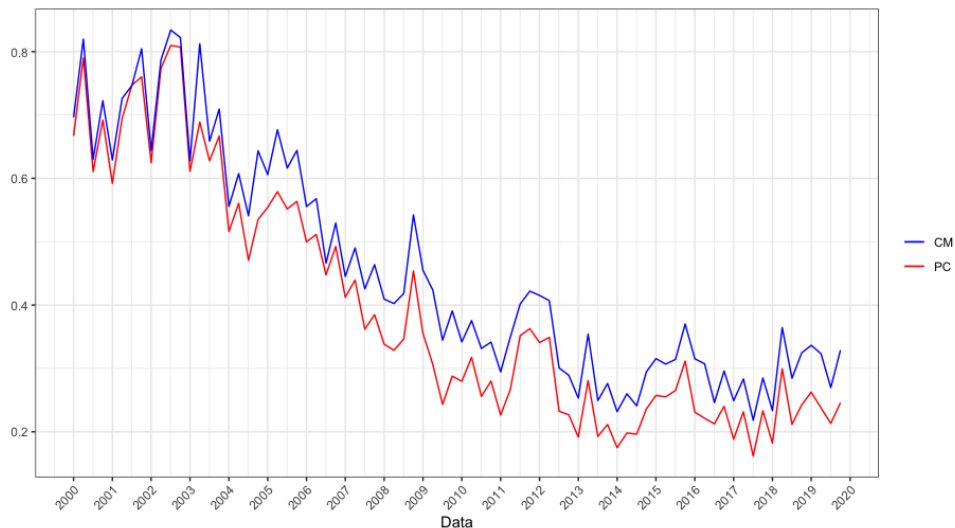
Na figura 16 observa-se que tanto a média por período do Custo Médio ($\overline{CM}$) quanto a média por período do Preço do Crédito ($\overline{PC}$) não são constantes e possuem valores que variam ao longo do tempo, logo, posição geográfica muito próximos⁴³, mas não similares. O $\overline{CM}$ possui valores pouco superiores ao $\overline{PC}$, sugerindo que, em relação ao produto crédito, as firmas da indústria bancária brasileira trabalham na média com lucro

⁴² A variável Receita da Intermediação Financeira é utilizada como *proxy* para a quantidade de crédito. Ver BANCO CENTRAL DO BRASIL (2018).

⁴³ Tabela 15 no Apêndice E.

negativo. Tal resultado colabora com as ponderações de Belaisch (2003), que sustenta que as grandes firmas tentem a minimizar seu risco de exposição na intermediação financeira ao reduzir sua participação no mercado de crédito de varejo priorizando ofertar crédito para o governo. Também colabora para esta sugestão, o elevado grau de inadimplência da economia brasileira e as ponderações de Barbosa, Rocha e Salazar (2014) quanto aos múltiplos produtos ofertados pelas instituições bancárias. Assim sendo, presume-se que a lucratividade da indústria bancária no Brasil possua um vínculo maior com a rentabilidade obtida nos títulos do governo e com os demais produtos e serviços ofertados por essas instituições, do que com mercado de crédito de varejo. Observa-se também que a tendência dos dois gráficos é negativamente inclinada, sugerindo que com o passar dos trimestres, tanto o $\overline{PC}$ quanto o $\overline{CM}$, estão reduzindo. Este resultado pode sugerir uma alteração na relação crédito varejo-governo, uma vez que a taxa básica de juros paga pelo governo reduziu.

Figura 16 – Curva da *proxy* de Custo Médio e Preço do Crédito.



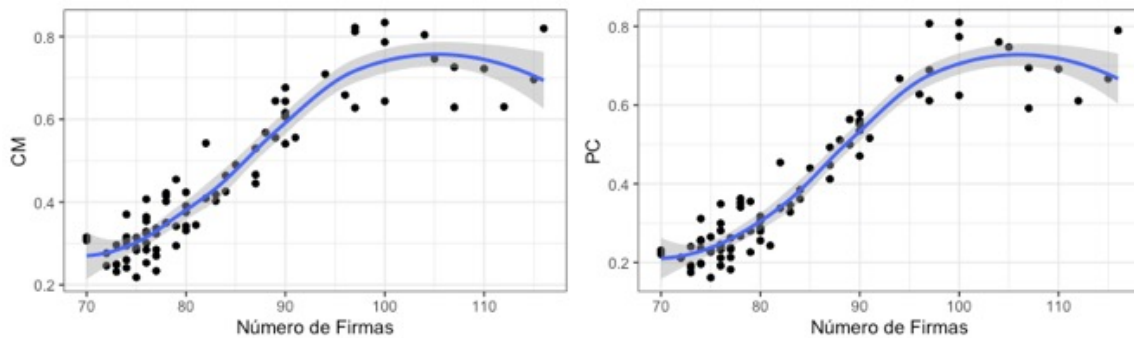
Fonte: Resultados da pesquisa.

Nota: No eixo Y são reportados os valores em log para a média trimestral da indústria das variáveis Preço do Crédito (PC) e Custo Médio (CM).

Os gráficos reportados na figura 17, que relacionam $\overline{CM}$ e $\overline{PC}$ – ambas no eixo y – ao Número de Firms por Trimestre – eixo x – comporta-se contrario à suposição inicial de que, quanto maior o mercado, maior a concorrência, menor os preços e menor os custos médios. O que se verifica nos gráficos é uma concentração de pontos correspondentes a menor número de firmas com menores valores das variáveis $\overline{CM}$ e $\overline{PC}$. É importante ressaltar alguns pontos que contribuem para a redução do número de firmas: i) redução da participação do estado; ii) fusões e aquisições e iii) saída de *players* em períodos de crise e devido a alterações nas políticas macroeconômicas do país. O comportamento observado nos dados contribui para suposição de que a digitalização bancária tem incentivado a

redução dos custos médio e do preço do crédito, haja vista que o número de instituições bancária tem uma ascendente queda entre os anos de 2000 até 2016, quando há uma leve reversão deste quadro, e é observado o crescimento do número de firmas na indústria.

Figura 17 – Gráfico de dispersão das médias trimestrais do custo médio e *proxy* do preço do crédito.



Fonte: Resultados da pesquisa.

Nota: CM – Custo Médio e PC – Preço do Crédito. Valores em log. A linha azul é calculada via Mínimos Quadrados Ordinários, o sombreado cinza que envolve a linha azul representa o intervalo de confiança de 95%.

Logo, averigua-se que, com o passar dos anos, o número de firmas operando no setor bancário brasileiro reduziu, entretanto, os custos médios e os preços do crédito também diminuiriam. Quanto a variação no número de empresas, constata-se um grande volume de fusões e aquisições, que possibilitaram as empresas a utilizar de forma mais eficiente as economias de escala. Contudo, também é observado a entrada e saída de firmas do setor. Tal variação no número de empresas serve de amparo a hipótese de inexistência de barreiras. O valor do $\overline{CM}$ é próximo ao valor do $\overline{PC}$, neste sentido, supõe-se que o Custo Marginal também possua um valor próximo, mas não igual ao do $\overline{PC}$ e o Preço do Crédito varia de firma para firma. Isto posto, a hipótese de que o mercado bancário brasileiro atua em um regime com características de competição monopolística é fortalecido.

4.4 ESTATÍSTICA H DE PANZAR-ROSSE

Na literatura consultada para este ensaio é evidente o uso de diversas variáveis dependentes para o cálculo da estatística H de Panzar-Rosse – coeficiente β da regressão. Dentre as quais, na bibliografia voltada a indústria bancária brasileira, se destaca o uso da Receita Total (modelo original de Panzar-Rosse), *Return on Assets* considerado em muitos estudos a melhor variável para se observar a estatística H no longo prazo e da Receita da Intermediação Financeira. Neste sentido, a tabela 10 reporta o resultado da

regressão da equação da receita na forma reduzida⁴⁴ das três variáveis dependentes – RIF, RT e ROA – nos modelos *Pooled*, Efeito Fixo e Efeito Aleatório. Doravante as especificações recebem a nomenclatura H^{RIF} para o modelo no qual a variável dependente é a Receita da Intermediação financeira, H^{RT} quando a variável dependente é a Receita Total e H^{ROA} para o modelo com a variável *Return on Assets*.

⁴⁴ O modelo da estatística H de Panzar-Rosse deriva de um sistema de equações interdependentes que é reduzido uma vez que utiliza na composição da variável custo somente fatores de produção exógenos a firma. Exclui-se por exemplo os custos operacionais (variável endógena a firma). Assim, na equação, tudo que está à direita é exógeno ao sistema e a esquerda endógeno.

Tabela 10 – Regressão - Pannel, Equação da receita na forma reduzida.

	<i>Variável dependente:</i>								
	RIF			RT			ROA		
	(Pooled)	(Ef. Fixo)	(Ef. Aleatório)	(Pooled)	(Ef. Fixo)	(Ef. Aleatório)	(Pooled)	(Ef. Fixo)	(Ef. Aleatório)
DT	0.036*** (0.013)	0.035*** (0.013)	0.036*** (0.013)	−0.007 (0.009)	−0.008 (0.010)	−0.007 (0.009)	0.001 (0.003)	0.0002 (0.002)	0.001 (0.003)
w	0.824*** (0.018)	0.819*** (0.018)	0.824*** (0.018)	0.786*** (0.013)	0.782*** (0.014)	0.786*** (0.013)	0.003 (0.004)	0.001 (0.003)	0.003 (0.004)
CRD	−0.044* (0.025)	−0.047* (0.025)	−0.044* (0.025)	−0.082*** (0.019)	−0.083*** (0.019)	−0.082*** (0.019)	−0.012** (0.005)	−0.013*** (0.005)	−0.012** (0.005)
TB	−0.024 (0.045)	−0.028 (0.054)	−0.024 (0.045)	0.005 (0.034)	−0.011 (0.040)	0.005 (0.034)	0.0005 (0.009)	0.0005 (0.010)	0.002 (0.010)
Constante	0.001 (0.005)		0.001 (0.005)	0.003 (0.004)		0.003 (0.004)	−0.002** (0.001)		−0.004** (0.002)
Observações	6,582	6,582	6,582	6,582	6,582	6,582	6,582	6,582	6,582
R ²	0.252	0.247	0.252	0.349	0.344	0.349	0.001	0.001	0.002
R ² <i>Ajustado</i>	0.252	0.228	0.252	0.349	0.327	0.349	0.001	−0.024	0.001
F Statistic	553.864*** (df = 4; 6577)	527.770*** (df = 4; 6419)	2,215.454***	881.985*** (df = 4; 6577)	840.773*** (df = 4; 6419)	3,527.940***	2.025* (df = 4; 6577)	2.301* (df = 4; 6419)	8.385*

Nota: *p<0.1; **p<0.05; ***p<0.01

Fonte: Resultados da pesquisa.

Nota: RIF – Receita da Intermediação Financeira, RT – Receita Total e ROA – *Return on Assets*.

Através da análise do coeficiente dos Insumos Exógenos (w) – a estatística H de Panzar-Rosse – verifica-se, que quando a variável dependente é a Receita da Intermediação Financeira (RIF) ou a Receita Total (RT), os coeficientes são estatisticamente significante ao nível de 1% nos três modelos – *Pooled*, Efeitos Fixos e Efeitos Aleatório. Todavia, quando a variável dependente é o *Return on Assets* (ROA), nenhum dos três modelos apresenta significância estatística. Ademais, na especificação empírica H^{ROA} , somente a variável Risco da Intermediação Financeira (CRD) apresenta coeficientes estatisticamente significativos – a 5% nos modelos *Pooled* e de Efeitos Aleatórios e 1% no modelo de efeito fixo. Observa-se também que, para as estimativas utilizando a variável dependente ROA, o poder de explicação dos modelos, medidos pelos R^2 e R^2 ajustado, são próximos de zero. Vale lembrar também que as correlações da variável ROA, apresentadas na tabela 7, com as demais variáveis do modelo – DT, w , CRD e TB – também são próximas de zero, – 0, – 0,03, 0,03 e 0 respectivamente. Assim como no trabalho de Silveira (2017) o β da variável w na especificação H^{ROA} não apresenta resultados significativos, neste sentido, a estimativa é descartada das demais análises realizadas neste ensaio.

Ainda na tabela 10 observa-se que o coeficiente da variável DT na especificação H^{RIF} possui sinal positivo e é estatisticamente significativo nos três modelos. Tal resultado é condizente com a literatura, uma vez que o crescimento da receita da intermediação financeira está diretamente relacionando ao crescimento dos depósitos. Porém, na especificação H^{RT} , a variável DT apresenta sinal negativo nos três modelos. Tal resultado poderia servir como base para o argumento de que a receita total dos bancos brasileiros na média está muito mais vinculada aos multiprodutos e serviços do que com a oferta de crédito no varejo, contudo, a variável não apresenta significância estatística. Já a variável CRD apresenta sinal negativo nas duas especificações H^{RIF} e H^{RT} , com significância estatística de 1% nos três modelos para a especificação cuja variável dependente é RT. E significância estatística de 10% na especificação com variável dependente RIF. O coeficiente da variável TB não apresenta significância estatística e apresenta sinal positivo nos modelos *Pooled* e de Efeitos Aleatórios da especificação H^{RT} .

4.4.1 Testes dos modelos

No intuito de alavancar a qualidade da estimativa H , são realizados testes para verificar qual modelo é mais eficiente. Através do teste F de *Chow*, inicialmente compara-se o modelo *Pooled* com a regressão de Efeitos Fixos. A hipótese nula pressupõe igualdade nos interceptos e nas inclinações para todos os indivíduos, caracterizando o *Pooled* como mais adequado. O p-valor é igual a 1 para nas especificações H^{RIF} e H^{RT} , logo aceita-se H_0 , o modelo *Pooled* é mais adequado que o modelo de Efeitos Fixos.

Em seguida foi realizado o teste do multiplicador de *Lagrange Breush-Pagan* (Modelo *Pooled* x Modelo de Efeitos Aleatórios). O teste compara as estimativas entre os modelos, verificando a hipótese nula – $\sigma_\alpha^2 = 0$ – e alternativa – $\sigma_\alpha^2 \neq 0$. Não rejeitar H_0 implica que o modelo de dados agrupados é mais adequado. O teste reporta p-valores menores que 1 (p valor de H^{RIF} é 7.991e-10 e o p-valor de H^{RT} é 3.772e-10). Assim H_0 deve ser rejeitada indicando que o modelo de Efeitos Aleatórios é mais adequado que o modelo *Pooled*.

Por fim, é feito o teste de *Hausman* para verificar se existe diferença estatisticamente significativa entre os parâmetros do modelo de Efeitos Fixo e Aleatório. Diferenças significativas entre elas indicam a inconsistência dos estimadores de Efeitos Aleatórios. Rejeitar H_0 constitui afirmar que existe correlação entre os efeitos e os regressores e, logo, os estimadores do modelo de Efeitos Fixos são consistentes. Os p-valores estimados são 0,6325 e 0,5822 para H^{RIF} e H^{RT} , portanto, a hipótese nula não deve ser rejeitada em face de H_A .

Desta forma, tem-se que as especificações H^{RIF} e H^{RT} apontam para o modelo de Efeitos Aleatórios como o mais adequado.

Dependência Transversal dos modelos de efeitos aleatório

A dependência transversal (*cross-sectional*) surge em painéis com longas séries temporais. Neste sentido, no teste de *Pesaran* H_0 pressupõe que os resíduos não estão correlacionados com os indivíduos. Os testes reportam p-valor $< 2.2e - 16$ para ambas as especificações H^{RIF} e H^{RT} , desta forma, rejeita-se H_0 , pois não é verificado dependência *cross-sectional*.

Homocedasticidade dos resíduos dos modelos de efeitos aleatório

O teste *Studentized Breusch-Pagan* verifica a existência ou não de variância constante (homocedasticidade) dos resíduos. A hipótese nula implica homocedasticidade, ou seja, pressupõe-se que a dispersão dos resíduos é constante. Os testes reportam p-valor igual a 0,0399 para o modelo com a variável dependente RIF e p-valor $< 2.467e - 05$ para a especificação H^{RT} . Os p-valores estimados indicam que ao nível de significância de 5% as especificações H^{RIF} e H^{RT} apresentam problemas de heterocedasticidade. Contudo, ao nível de significância de 1%, os resíduos da especificação H^{RIF} tendem a homocedas-

ticidade.

Correlação serial dos modelos de efeitos aleatório

Ademais, o teste de correlação serial – *Breusch-Godfrey/Wooldridge* – cuja hipótese nula é a de que não se encontre esta característica na série, também resultou em p-valor $< 2.2e - 16$ para H^{RT} , e $< 3.555e - 15$ para H^{RIF} , designando correlação serial nos erros idiossincráticos.

Variância robusta

Dados os resultados obtidos nos teste de *Pesaran e Breusch-Godfrey/Wooldridge*, é realizado a estimação dos modelos de Efeitos Aleatórios para as especificações cuja variável dependente são a RIF e RT por variância robusta (*White e Arellano*).

Dos resultados reportados na tabela 10 para as especificações H^{RIF} e H^{RT} no modelo de efeitos aleatórios, observa-se que o coeficiente β da variável Insumos Exógenos (w) apresentam sinal positivo e coeficiente estatisticamente significativo a 1% (0,824 e 0,786 respectivamente). O coeficiente não perde significância estatística na estimação pelo método *White*, contudo, no método *Arellano* verifica-se não só a perda de significância estatística como a inversão do sinal e mudança brusca nos valores dos coeficientes. As variáveis DT e CRD da especificação H^{RIF} perdem toda a significância estatística nos dois métodos. Na especificação H^{RT} a variável CRD mantém a significância estatística observada na regressão pelo método *White*, mas perde no *Arellano*. Destaca-se que a variável TB que não apresenta significância estatística em nenhum modelo da tabela 10, porém, no modelo *Arellano* é estatisticamente significante a 1%.

Variância robusta com e sem os bancos digitais

A análise a seguir mostra o resultado das regressões no modelo de Efeitos Aleatórios utilizando o método de variância robusta de *White* em amostras com e sem os bancos digitais. O número de observações reduz em 70 unidades nas especificações sem os bancos digitais. Outro ponto que merece destaque é a exclusão da variável *dummy* Tipo Bancário

Tabela 11 – Variância Robusta das especificações H^{RIF} e H^{RT}

<i>Variável dependente:</i>				
	(W-RIF)	(W-RT)	(A-RIF)	(A-RT)
DT	0.0359 (0.0291)	-0.0074 (0.0190)	0.0065 (0.0214)	0.0065 (0.0189)
w	0.8241*** (0.0353)	0.7860*** (0.0299)	-0.0137 (0.0390)	-0.0137 (0.0387)
CRD	-0.0438 (0.0445)	-0.0817*** (0.0270)	0.0196 (0.0240)	0.0196 (0.0294)
TB	-0.0241 (0.0192)	0.0050 (0.0377)	-0.1222*** (0.0097)	-0.1222*** (0.0107)
Constante	0.0014 (0.0048)	0.0027 (0.0036)		

Nota:

*p<0.1; **p<0.05; ***p<0.01
Fonte: Resultados da pesquisa.

Nota: W-RIF – método *White* com variável Receita da Intermediação Financeira, W-RT – método *White* com variável Receita Total, A-RIF – método *Aellano* com variável Receita da Intermediação Financeira e A-RT – método *Arellano* com variável Receita Total.

(TB) no modelo sem os bancos digitais.

Verifica-se na tabela 12 que entre as especificações com os bancos digitais e sem os bancos digitais não há perda de significância estatística e alteração dos sinais. O coeficiente da variável w, que representa a estatística H de Panzar-Rosse, é maior na especificação H^{RIF} com os bancos digitais. Porém, menor na especificação H^{RT} . Ao extrair os bancos digitais do modelo, a variável DT cresce na especificação H^{RIF} e diminui na H^{RT} . Já a variável CRD é maior na especificação H^{RIF} sem os bancos digitais e maior na H^{RT} com os bancos digitais inclusos na regressão.

Constata-se que os valores estimados da estatística H – coeficiente da variável w – rejeitam a hipótese de que o setor bancário brasileiro é monopolista, pois não satisfazem o teorema 1, que define que a soma das elasticidades-preço dos insumos exógenos (w) da equação de receita na forma reduzida de um monopolista devem ser menores que zero. Também não é satisfeita a proposição 2, que declara que para firmas observadas em equilíbrio competitivo de longo prazo, a soma das elasticidades da receita da firma na forma reduzida em relação ao preço dos fatores é igual a um. Logo, rejeita-se a hipótese de mercado em equilíbrio competitivo de longo prazo. Entretanto, os resultados obtidos

Tabela 12 – Variância Robusta

	<i>Variável dependente:</i>			
	(W-RIF)	(W-RT)	(W-RIF-STB)	(W-RT-STB)
DT	0.0359 (0.0291)	-0.0074 (0.0190)	0.0370 (0.0293)	-0.0097 (0.0187)
w	0.8241*** (0.0353)	0.7860*** (0.0299)	0.8233*** (0.0355)	0.7881*** (0.0301)
CRD	-0.0438 (0.0445)	-0.0817*** (0.0270)	-0.0433 (0.0447)	-0.0836*** (0.0269)
TB	-0.0241 (0.0192)	0.0050 (0.0377)		
Constante	0.0014 (0.0048)	0.0027 (0.0036)	0.0014 (0.0048)	0.0027 (0.0036)
Observações	6582	6582	6512	6512
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Fonte: Resultados da pesquisa.

Nota: W-RIF – método *White* com variável Receita da Intermediação Financeira, W-RT – método *White* com variável Receita Total, W-RIF-STB – método *White* com variável Receita da Intermediação Financeira sem os bancos digitais, W-RT-STB – método *White* com variável Receita Total sem os bancos digitais.

indicam que o mercado bancário brasileiro é condizente com o equilíbrio simétrico de *Chamberlin*, pois a soma das elasticidades da receita da firma na forma reduzida, em relação ao preço dos fatores (insumos exógenos), são menores que um ($White H_{RIF} = 0.8241$, $White H_{RT} = 0.7860$, $White H_{RIF}$ sem bancos digitais = 0.8233 e $White H_{RT}$ sem bancos digitais = 0.7881).

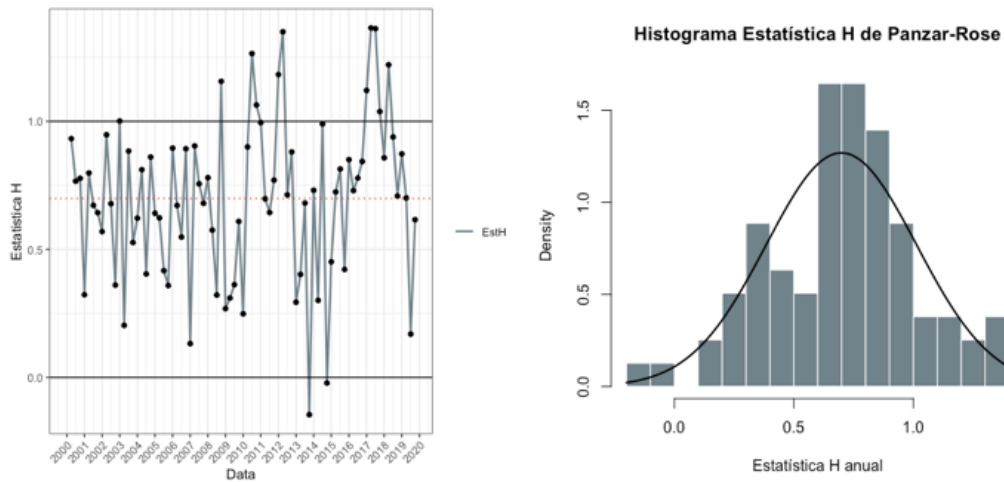
Para o objetivo da pesquisa – analisar o impacto do crescimento dos bancos digitais sobre a estrutura de mercado da indústria bancária brasileira sob a ótica da concentração, utilizando o Índice de *Herfindahl-Hirschman* normalizado, e da competição, empregando a metodologia da estatística H – os resultados estimados com a técnica de Panzar-Rosse não são significativos na competição do setor. Primeiro, porque a variável Tipo Bancário (TB) não é estatisticamente significativa e reporta sinal negativo na especificação $White H_{RIF}$. Segundo, porque não é observado nenhuma variação considerável nos valores do βW obtidos nas especificações com e sem os bancos digitais.

4.4.2 Inferência Estatística H de Panzar-Rosse

No intuito de verificar um intervalo de confiança para a estatística H de Panzar-Rosse, são realizadas regressões trimestrais das especificações H^{RIF} e H^{RT} por mínimos quadrados ordinários da amostra dessazonalizada, em logaritmo e em diferença de primeira ordem.⁴⁵

Nas imagens a seguir estão representados os gráficos das estimativas da estatística H obtidas para o intervalo entre o segundo trimestre de 2000 até o quarto trimestre de 2019 e o histograma da distribuição calculada.

Figura 18 – Gráfico de dispersão e histograma da estimativa da estatística H^{RIF} de Panzar-Rosse entre o segundo trimestre de 2000 até o quarto trimestre de 2019.

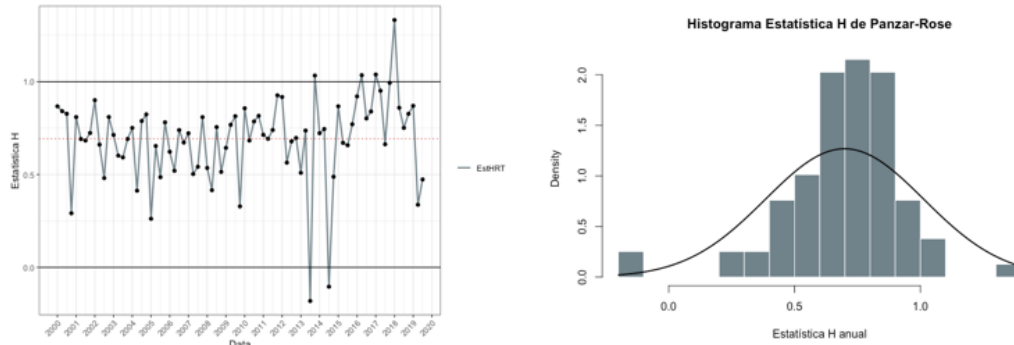


Fonte: Resultados da pesquisa.

Nota: As linhas horizontais contínuas pretas representam os intervalos em que $H^{RIF} \leq 0$ indica mercado com características de monopólio, $H^{RIF} \geq 1$ indica mercado com características de concorrência perfeita. A linha horizontal tracejada de vermelho marca a média da distribuição.

⁴⁵ Os resultados estão reportados na tabela 15 no apêndice E.

Figura 19 – Gráfico de dispersão e histograma da estimativa da estatística H^{RT} de Panzar-Rosse entre o segundo trimestre de 2000 até o quarto trimestre de 2019.



Fonte: Resultados da pesquisa.

Nota: As linhas horizontais contínuas pretas representam os intervalos em que $H^{RT} \leq 0$ indica mercado com características de monopólio, $H^{RT} \geq 1$ indica mercado com características de concorrência perfeita. A linha horizontal tracejada de vermelho marca a média da distribuição.

Observa-se que as duas variáveis são bem dispersas no tempo, com alguns valores que ultrapassam as barreiras do monopólio ($H \leq 0$) e da concorrência perfeita ($H \geq 1$). Contudo, também se verifica que a maior parte das estimativas estão próximas da média. No histograma das figuras 18 e 19, verifica-se que as estimativas tendem a normalidade.

Para verificar o intervalo de confiança das estimativas são calculadas as estatísticas t e Z separadamente para os níveis de confiança de 90, 95 e 99%.

Tabela 13 – Intervalo de confiança da distribuição trimestral das estimações das estatísticas H^{RIF} e H^{RT} entre o segundo trimestre de 2000 até o quarto trimestre de 2019

Variável	Nível de confiança	Teste	Intervalo de confiança	
H^{RIF}	90%	t	0.6399	0.7577
H^{RT}	90%	t	0.6501	0.7357
H^{RIF}	90%	z	0.6565	0.7410
H^{RT}	90%	z	0.6506	0.7351
H^{RIF}	95%	t	0.6283	0.7692
H^{RT}	95%	t	0.6417	0.7440
H^{RIF}	95%	z	0.6484	0.7491
H^{RT}	95%	z	0.6425	0.7432
H^{RIF}	99%	t	0.6054	0.7922
H^{RT}	99%	t	0.6251	0.7607
H^{RIF}	99%	z	0.6326	0.7649
H^{RT}	99%	z	0.6267	0.7591
H^{RIF}	Média	0.6988	Desvio Padrão	0.3144
H^{RT}	Média	0.6929	Desvio Padrão	0.2284

Fonte: Resultados da pesquisa.

Nota: H^{RIF} – especificação no qual a variável dependente é a Receita da Intermediação Financeira, H^{RT} – especificação no qual a variável dependente é a Receita Total.

Dois pontos precisam ser lembrados ao se analisar os resultados estimados para os intervalos de confiança. Primeiro: o modelo que melhor se ajusta para os dados em painel é o modelo de Efeitos Aleatórios e a estimação por trimestre não comporta este modelo. Segundo: diante dos resultados dos testes de *Pesaran* e *Breush-Godfrey/Wooldridge*, a estimativa dos dados em painel foi calculada utilizando o método de *White*, a estimação por trimestre não. Observado estes aspectos, constata-se que nenhum resultado obtido nas estimações da estatística H – método *White* especificação H^{RIF} com e sem bancos digitais e método *White* especificação H^{RT} com e sem bancos digitais – encontra-se dentro dos intervalos de confiança calculados pelos testes t e Z . Contudo, observa-se na tabela 13 que todos os valores reportados nos intervalos de confiança encontram-se no intervalo de interpretação da estatística H configurada como concorrência monopolística. O mesmo resultado se observa na tabela 12, os quatro coeficientes da variável w são significativos a 1% e são números positivos menores que um. Neste sentido, verifica-se neste ensaio, que o modelo simplificado, como preconizado por Panzar e Rosse (1987), entregou, no âmbito nacional, resultados condizentes com Rondon (2011), Neto, Araújo e Ponce (2006) e grande parte da literatura consultada. Implicando em uma indústria bancária condizente com concorrência monopolística.

5 CONSIDERAÇÕES FINAIS

Neste estudo busca-se analisar o impacto do crescimento dos bancos digitais sobre a estrutura de mercado da indústria bancária brasileira sob a ótica da concentração, utilizando o Índice de *Herfindahl-Hirschman* normalizado, e da competição, empregando a metodologia da estatística H de Panzar-Rosse. Afere-se índices de *Herfindahl-Hirschman* normalizado próximos aos reportados pelo Banco Central do Brasil para as variáveis Depósito Total, Saldo das Operações de Crédito e Ativo Total Ajustado no mês de dezembro dos anos de 2017, 2018 e 2019, indicando uma indústria bancária com níveis moderados de concentração. Também se afere valores para a estatística H de Panzar-Rosse com valores variando entre 0,7860 e 0,8241, sugerindo um mercado com características de competição monopolística.

No que tange ao objetivo principal do ensaio, o índice de *Herfindahl-Hirschman* normalizado aponta que o crescimento dos bancos digitais na indústria tende a reduzir os níveis de concentração. O resultado apresentado para o IHH_n mostra que apesar de a diferença ser pequena, as três variáveis com os bancos digitais, apresentam níveis de concentração menores que as variáveis na qual os bancos digitais não estão presentes na indústria. O mesmo resultado não é observado com os resultados considerados na estatística H de Panzar-Rosse. Os dados não mostram que o crescimento do número de bancos digitais na indústria tem algum efeito significativo no nível de competição. Tais resultados podem estar relacionados com: os filtros utilizados, uma vez que na categoria b1 (Bancos Comercial, Banco Múltiplo com Carteira Comercial ou Caixas Econômicas) não estão enquadradas as Instituições de Pagamentos, o período da amostra, já que se observa que o número de bancos digitais está crescendo e atualmente já é maior do que no intervalo temporal coberto pelo estudo e a limitações da estatística H para o objetivo apresentado. Vale ressaltar também que não é só o número de Bancos digitais que está em ascendência, mas a adesão de clientes a este tipo de serviço, que tende a impulsionar as variáveis chave para o cálculo dos níveis de concorrência.

O estudo mostra-se eficiente ao elencar normativas adotadas pelo Banco Central do Brasil que visam promover a concorrência e desconcentrar o mercado, além de proporcionar um ambiente institucional propício para o crescimento da oferta de serviços digitais na indústria bancária. Tais medidas mostram-se fundamentais tendo em vista a constante redução do número de agências bancárias e a desigualdade na distribuição geográfica do serviço físico. O trabalho contribui para as proposições expostas por Belaisch (2003) quanto à ineficiência da indústria bancária no produto crédito de varejo – as firmas tentam a minimizar os riscos na intermediação financeira reduzindo sua participação no mercado de crédito de varejo e priorizam a oferta de crédito para o governo – e para as hipótese levantada por Barbosa, Rocha e Salazar (2014) quanto à estrutura multiproduto

do setor bancário ser forte diferencial entre bancos superavitários de deficitários. A metodologia proposta também revela-se competente por deixar uma herança replicável. Neste sentido, uma sugestão para a busca de resultados mais conclusivos para avaliar o grau de competição, caso haja disponibilidade de dados, seria a utilização dos Indicadores de *Lerner* e *Boone* que permitem avaliar a competição usando a razão entre o *mark-up* e o preço e estimativas do custo marginal, que podem fornecer medidas mais adequadas para analisar os níveis de competição em uma indústria.

REFERÊNCIAS

- BANCO CENTRAL DO BRASIL. Relatório de economia bancária. BACEN, Brasília, DF, 2018. 215 p. Disponível em: <<https://www.bcb.gov.br/publicacoes/relatorioeconomiabancaria/31122018>>. Acesso em: 16 mar. 2020.
- BANCO CENTRAL DO BRASIL. Estatística bancária mensal por município. BACEN, Brasília, DF, 2019. Disponível em: <<https://www4.bcb.gov.br/fis/cosif/estban.asp>>. Acesso em: 07 jul. 2020.
- BANCO CENTRAL DO BRASIL. Relatório de economia bancária. BACEN, Brasília, DF, 2019. 213 p. Disponível em: <https://www.bcb.gov.br/publicacoes/relatorioeconomiabancaria/REB_2019>. Acesso em: 16 mar. 2020.
- BANCO CENTRAL DO BRASIL. Relatórios trimestrais. BACEN, Brasília, DF, 2019. Disponível em: <<https://www3.bcb.gov.br/ifdata/>>. Acesso em: 05 mai. 2020.
- BANCO CENTRAL DO BRASIL. Relatório de economia bancária. BACEN, Brasília, DF, 2020. 253 p. Disponível em: <https://www.bcb.gov.br/publicacoes/relatorioeconomiabancaria/reb_2020>. Acesso em: 17 mar. 2020.
- BARBOSA, Klenio; ROCHA, Bruno de Paula; SALAZAR, Fernando. Assessing competition in the banking industry: A multi-product approach. *Journal of Banking & Finance*, v. 50, p. 340–362, 2014.
- BAUMOL, William J.; PANZAR, John C.; WILLIG, Robert D. Contestable markets: An uprising in the theory of industry structure: Reply. *The American Economic Review*, v. 73, n. 3, p. 491–496, 1983.
- BELAISCH, Ms Agnes. Do brazilian banks compete? *International Monetary Fund*, v. 2003, n. XX, p. 22, Jun 2003.
- BELLEFLAME, Paul; PEITZ, Martin. *Industrial organization*. Cambridge: Cambridge University Press, 2010. 726 p.
- BIKKER, Jacob A.; SHAFFERS, Sherrill; SPIERDIJK, Laura. Assessing competition with the panzar-rosse model: The role of scale, costs, and equilibrium. *Review of Economics and Statistics*, v. 94, n. 4, p. 1025–1044, 2012.
- BIKKER, Jacob A.; SPIERDIJK, Laura; FINNIE, Paul. Misspecification of the panzar-rosse model: Assessing competition in the banking industry. *DNB Working Papers*, v. 114, n. XX, p. 1–31, Out 2006.
- BRASIL. Banco Central do Brasil. Resolução nº 3.919, de 25 de novembro de 2013. altera e consolida as normas sobre cobrança de tarifas pela prestação de serviços por parte das instituições financeiras e demais instituições autorizadas a funcionar pelo banco central do brasil e dá outras providências. *DIÁRIO OFICIAL DA UNIÃO*, Brasília, DF, 04 nov. 2013.

BRASIL. Banco Central do Brasil. Resolução nº 4.283, de 4 de novembro de 2013. altera a resolução nº 3.694, de 26 de março de 2009, que dispõe sobre a prevenção de riscos na contratação de operações e na prestação de serviços por parte de instituições financeiras e demais instituições autorizadas a funcionar pelo banco central do Brasil. DIÁRIO OFICIAL DA UNIÃO, Brasília, DF, 04 nov. 2013. Revogado.

BRASIL. Banco Central do Brasil. Resolução nº 4.480, de 25 de abril de 2016. dispõe sobre a abertura e o encerramento de contas de depósitos por meio eletrônico e dá outras providências. DIÁRIO OFICIAL DA UNIÃO, Brasília, DF, 26 abr. 2016. Seção 1, Edição 78, Página 15.

BRASIL. Banco Central do Brasil. Resolução nº 4.630, de 25 de janeiro de 2018. altera a resolução nº 4.480, de 25 de abril de 2016, que dispõe sobre a abertura e o encerramento de contas de depósitos por meio eletrônico. DIÁRIO OFICIAL DA UNIÃO, Brasília, DF, 25 jan. 2018. Revogado.

BRASIL. Banco Central do Brasil. Resolução nº 4.639, de 22 de fevereiro de 2018. altera a resolução nº 3.402, de 6 de setembro de 2006, que dispõe sobre a prestação de serviços de pagamento de salários, aposentadorias e similares sem cobrança de tarifas. DIÁRIO OFICIAL DA UNIÃO, Brasília, DF, 30 abr. 2018. Seção 1, Edição 82, Página 24.

BRASIL. Banco Central do Brasil. Resolução nº 4.656, de 26 de abril de 2018. dispõe sobre a sociedade de crédito direto e a sociedade de empréstimo entre pessoas, disciplina a realização de operações de empréstimo e de financiamento entre pessoas por meio de plataforma eletrônica [...]. DIÁRIO OFICIAL DA UNIÃO, Brasília, DF, 30 abr. 2018. Seção 1, Edição 82, Página 24.

BRASIL. Banco Central do Brasil. Resolução nº 4.697, de 27 de novembro de 2018. altera a resolução nº 4.480, de 25 de abril de 2016, que dispõe sobre a abertura e o encerramento de contas de depósitos por meio eletrônico. DIÁRIO OFICIAL DA UNIÃO, Brasília, DF, 02 fev. 2018. Seção 1, Edição 38, Página 9–46.

CARDOSO, Marcelo; AZEVEDO, Paulo Furquim de; BARBOSA, Klênio de Souza. Concorrência no setor bancário brasileiro: Bancos individuais versus conglomerados bancários. *IPEA, Pesquisa e Planejamento Econômico (PPE)*, Rio de Janeiro, RJ, v. 46, n. 01, p. 113–146, 2016.

CORNWALL, Gary; CHEN, Jeff. The dark side of the moon: Searching for the other half of seasonality. *Bureau of Economic Analysis*, XX, n. XX, Nov 2019.

CROISSANT, Yves; MILLO, Giovanni. Panel data econometrics in r: The plm package. *Journal of statistical software*, v. 27, n. 2, 2008.

DE-LOSSO, Rodrigo. Deflacionamento. FIPE, XX, n. XX, jul 2020. Disponível em: <<https://downloads.fipe.org.br/publicacoes/bif/bif478-18-24.pdf>>. Acesso em: 15 Dez. 2020.

DIVINO, José Angelo; SILVA, Renan Said. Banking competition in the Brazilian economy. *Nova Economia*, v. 27, n. 3, p. 393–417, Sep-Dec 2017.

FERREIRA, Pedro GC.; JR, José Lisboa Gondin; MATTOS, Daiane M. de. X-13arima-seats com r: Um estudo de caso para a produção industrial brasileira. Instituto Brasileiro de Economia (FGV| IBRE), XX, n. XX, 2015.

FERREIRA, Pedro GC.; MATTOS de Daiane M. Usando o r para ensinar ajuste sazonal. Instituto Brasileiro de Economia (FGV| IBRE), 2016. Disponível em: <https://200.255.81.16/data/files/E0/D2/E5/EB/7FD335108745A9155C28C7A8/Ferreira_Mattos_2016.pdf>. Acesso em: 26 jan. 2021.

GREENE, William H. *Econometric analysis*. 4. ed. New Jersey: Prentice Hall, 2000.

GUJARATI, Damodar N.; PORTER, Dawn C. *Econometria básica*. 5. ed. [S.l.]: Amgh Editora, 2011.

HORDONES, Cristiano; SANVICENTE, Antonio Zoratto. Estrutura, poder de mercado e rentabilidade: evidências do setor bancário na américa latina. *Revista Contabilidade & Finanças*, v. 32, p. 126–142, 2020.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Estimativas da população. IBGE, 2019. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?edicao=25272&t=resultados>>. Acesso em: 01 dez. 2020.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Shapefile. IBGE, 2019. Disponível em: <<https://www.ibge.gov.br/geociencias/downloads-geociencias.html>>. Acesso em: 10 out. 2020.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Índice nacional de preço ao consumidor amplo. IPEADATA, 2019. Disponível em: <<http://www.ipeadata.gov.br/Default.aspx>>. Acesso em: 01 dez. 2020.

JUSTUS, Marcelo; KASSOUF, Ana Lúcia. Uma investigação econômica da influência do mercado de drogas ilícitas sobre a criminalidade brasileira. *Revista Economia*, v. 8, n. 2, p. 187–210, 2007.

KADIR, Hazlina Abd et al. Market structure and bank competition conditions in malaysia. *Journal of Technology Management and Business*, v. 1, n. 2, 2014.

KRUGMAN, Paul R.; OBSTFELD, Maurice. *Economia Internacional – Teoria e Política*. 5. ed. São Paulo, SP: Makron Books, 2001.

LIPTON, Alex; SHRIER, David; PENTLAND, Alex. *Digital banking manifesto: the end of banks?* Massachusetts Institute of Technology, 2016. Disponível em: <https://www.getsmarter.com/blog/wp-content/uploads/2017/07/mit_digital_bank_manifesto_report.pdf>. Acesso em: 01 Abr. 2020.

LUCINDA, Claudio R. Competition in the brazilian loan market: An empirical analysis. *Estudos Econômicos (São Paulo)*, v. 40, n. 4, p. 831–858, 2010.

NETO, Paulo de Melo Jorge; ARAÚJO, Luiz Alberto D'Avila de; PONCE, David Agustín Salazar. Competição e concentração entre os bancos brasileiros. *Revista Economia*, 2006. Disponível em: <https://anpec.org.br/revista/vol7/vol7n3p561_586.pdf>. Acesso em: 16 Mar. 2021.

PANZAR, John C.; ROSSE, Janes N. Testing for "monopoly"equilibrium. *The journal of industrial economics*, XXXV, n. 4, p. 443–456, Jun 1987.

RONDON, Leonardo Velasco. *Competitividade e eficiência do sistema financeiro nacional: 1995-2008*. 221 f. Tese (Doutorado em Economia) — Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 2011. Disponível em: <<http://hdl.handle.net/1843/AMSA-8P2JGT>>. Acesso em: 15 jan. 2020.

SANTOS, Marcelo Henrique Shinkoda. *Conduta bancária nos segmentos de empréstimos e depósitos no brasil*. 168 f. Dissertação (Mestrado em Economia Aplicada) — Departamento de Economia, Universidade Federal de Viçosa, Viçosa, MG, 2018. Disponível em: <<https://locus.ufv.br/handle/123456789/21290>>. Acesso em: 07 jul. 2020.

SILVEIRA, Pedro Paulo Bartolomei da. *Concentração e competição bancárias no Brasil, uma aplicação do modelo Panzar-Rosse*. 48 p. Dissertação (Mestrado Profissional em Finanças e Economia) — Escola de Economia de São Paulo, Fundação Getúlio Vargas, São Paulo, 2017. Disponível em: <<http://hdl.handle.net/10438/18820>>. Acesso em: 15 jan. 2020.

SPIERDIJK, Laura; SHAFFER, Sherrill. The panzar-rosse revenue test and market power in banking. *Journal of Banking & Finance*, v. 61, p. 340–347, 2015.

TABAK, Benjamin M.; GOMES, Guilherme MR.; JR, Maurício da Silva Medeiros. The impact of market power at bank level in risk-taking: The brazilian case. *International Review of Financial Analysis*, v. 40, p. 154–165, 2015.

WOLF, Alyson Prado. *Evolução estrutural do mercado brasileiro de produtos e serviços financeiros do ano 2008 a 2018 e avaliação da maturidade dos novos entrantes*. 104 f. Dissertação (Mestrado em Desenvolvimento Econômico) — Departamento de Economia, Universidade Federal do Paraná, Curitiba PR, 2020. Disponível em: <<https://hdl.handle.net/1884/69101>>. Acesso em: 05 mai. 2021.

WOLF, Diogo. *A evolução da competição no setor bancário no brasil utilizando o modelo de Panzar Rosse*. 56 f. Monografia (Graduação em Ciências Econômicas) — Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre. RS, 2020. Disponível em: <<http://hdl.handle.net/10183/216268>>. Acesso em: 05 mai. 2021.

APÊNDICE A – TABELA DOS TESTES DE SAZONALIDADE

Tabela 14 – Variáveis que apresentaram sazonalidade

(continua)

CDGO	OBS	RIF	RT	ROA	DT	W	CRD	PC	CM	DTT	SOC	ATA
208	3											
10021	3											
10045	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
10069	80	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	SIM			
10076	34	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
10083	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
10100	35	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
20066	3											
20121	15											
20145	8											
20152	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
20231	2											
20255	64	NaN	SIM	NÃO	SIM	SIM	NÃO	NÃO	NÃO			
30135	4											
30142	2											
30159	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
30173	80	SIM	SIM	SIM	SIM	SIM	NÃO	SIM	SIM			
30207	68	SIM	SIM	SIM	NÃO	SIM	NÃO	NÃO	NÃO			
30245	10											
30276	30	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
30290	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
30300	26	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
30379	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
30403	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
30513	17											
30771	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
30881	52	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
31158	23	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
31165	29	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	SIM			
31244	54	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
31323	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
31330	15											
31677	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
31873	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
31976	77	SIM	SIM	SIM	SIM	SIM	SIM	SIM	NÃO			
32016	46	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
32078	14											
32119	64	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
41856	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
49906	80	SIM	SIM	SIM	SIM	SIM	NÃO	SIM	SIM			
49920	2											
49944	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
49951	54	SIM	SIM	SIM	SIM	SIM	NÃO	SIM	SIM			
50012	11											
50071	79	SIM	SIM	SIM	SIM	SIM	NÃO	SIM	SIM			
50122	61	SIM	SIM	SIM	NÃO	SIM	NÃO	NÃO	NÃO			
50146	9											
50177	19											
50191	46	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
50201	80	SIM	SIM	NÃO	SIM	SIM	NÃO	NÃO	NÃO			
50304	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
50328	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			

(continua)

CDGO	OBS	RIF	RT	ROA	DT	W	CRD	PC	CM	DTT	SOC	ATA
50414	32	SIM	SIM	NAO	NAO	SIM	NAO	SIM	NAO			
50452	13											
50476	54	SIM	SIM	SIM	NAO	SIM	NAO	SIM	NAO			
50524	35	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
50531	80	SIM	SIM	SIM	NAO	SIM	SIM	SIM	SIM			
50720	49	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
50861	38	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
50885	51	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
50940	11											
50988	80	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
51011	80	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
51059	29	NAO	SIM	NAO	NAO	SIM	NAO	NAO	SIM			
51066	80	NaN	SIM	NAO	NAO	SIM	NAO	NAO	SIM			
51097	8											
51114	46	NaN	SIM	NaN	NAO	SIM	NAO	SIM	NAO			
51121	32	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
51145	25	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
51152	66	SIM	SIM	NAO	SIM	SIM	NAO	SIM	SIM			
51183	80	NaNs	SIM	SIM	NAO	SIM	NAO	NaNs	SIM			
51217	37	NAO	NAO	NAO	NAO	SIM	NAO	NAO	NAO			
51255	35	SIM	SIM	SIM	NAO	SIM	NAO	NAO	NAO			
51293	80	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
51468	80	SIM	SIM	SIM	NAO	SIM	NAO	NAO	NAO			
51516	60	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
51602	8											
51705	9											
51736	34	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
51750	36	SIM	SIM	SIM	SIM	SIM	NAO	SIM	SIM			
51781	80	SIM	SIM	SIM	SIM	SIM	NAO	SIM	SIM			
51815	14											
51839	28	SIM	SIM	SIM	NAO	SIM	NAO	NAO	NAO			
51884	46	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
51901	9											
51949	20	SIM	SIM	NAO	NAO	SIM	NAO	NAO	NAO			
51956	34	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
51963	17											
51987	80	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
52003	12											
52010	11											
52034	80	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
52072	4											
253448	73	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
360305	80	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
517645	80	SIM	SIM	SIM	NAO	SIM	SIM	SIM	SIM			
558456	28	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
795423	66	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
1023570	41	NAO	SIM	SIM	NAO	SIM	SIM	SIM	SIM			
1181521	80	SIM	SIM	SIM	SIM	SIM	NAO	SIM	SIM			
1522368	20	NAO	NAO	NAO	NAO	SIM	NAO	SIM	SIM			
1540541	8											
2038232	44	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
2318507	80	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
3532415	17											
4184779	35	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
4562120	8											
4866275	9											
4902979	80	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
4913711	80	SIM	SIM	SIM	SIM	SIM	SIM	SIM	SIM			
6271464	16											
6833131	35	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
7237373	80	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
7679404	43	SIM	SIM	SIM	NAO	SIM	NAO	SIM	SIM			
9093352	7											
9391857	30	SIM	SIM	NAO	NAO	SIM	NAO	SIM	SIM			
9516419	12											
10690848	14											
11476673	22	SIM	SIM	NAO	NAO	SIM	NAO	NAO	NAO			

(conclusão)

CDGO	OBS	RIF	RT	ROA	DT	W	CRD	PC	CM	DTT	SOC	ATA
13009717	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
14388334	28	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
15173776	80	SIM	SIM	SIM	SIM	SIM	SIM	SIM	SIM			
15357060	29	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	NÃO			
17351180	80	SIM	SIM	SIM	SIM	SIM	NÃO	SIM	SIM			
17453575	23	SIM	SIM	NÃO	SIM	SIM	NÃO	SIM	SIM			
23522214	13											
28157204	6											
29030467	14											
30723886	63	SIM	SIM	SIM	NÃO	SIM	NÃO	NÃO	NÃO			
31880826	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
32254138	51	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
33042151	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	NÃO	NÃO			
33098518	5											
33132044	80	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
33436486	8											
33466988	12											
33588252	9											
40429946	6											
42593459	2											
43073394	36	SIM	SIM	SIM	SIM	SIM	SIM	SIM	SIM			
43717511	45	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
48795256	47	SIM	SIM	NÃO	NÃO	SIM	NÃO	NÃO	NÃO			
50290345	20	SIM	SIM	NaNs	NÃO	SIM	NÃO	SIM	NÃO			
55230916	18											
58616418	24	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
59118133	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
60498557	64	SIM	SIM	NÃO	NÃO	NaNs	NÃO	NÃO	SIM			
60518222	80	NaNs	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
61033106	25	NÃO	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
61182408	11											
61535100	16											
62421979	42	SIM	SIM	SIM	NÃO	SIM	NÃO	SIM	SIM			
69057453	4											
74828799	80	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
78336633	4											
83876003	24	SIM	SIM	NÃO	NÃO	SIM	NÃO	SIM	SIM			
92791813	5											
92864131	11											
92874270	71	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	SIM			

Fonte: Resultado da pesquisa.

Nota: Se a célula da tabela possui: VAZIO, significam que o código não foi testado para sazonalidade, dado o baixo número de observações. SIM, significa que a variável apresentou sazonalidade, e foi feito o ajuste na série temporal. NÃO, significa que a variável não apresentou sazonalidade, considerando o nível de significância de 95%.

APÊNDICE B – DIAGNÓSTICO DO PRÉ-AJUSTE E O MODELO ARIMA

Diagnóstico do pré-ajuste e o modelo ARIMA da variável Receita da Intermediação Financeira (RIF) do banco 10045 no intervalo entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.

Call:

```
seas(x = RIF)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
Constant	-0.003087	0.002229	-1.385	0.166	
A02002.3	0.377824	0.039963	9.454	< 2e-16	\$~{***}\$
AR-Nonseasonal-01	-0.691968	0.111375	-6.213	5.2e-10	\$~{***}\$
AR-Nonseasonal-02	-0.019772	0.130970	-0.151	0.880	
AR-Nonseasonal-03	-0.008999	0.113286	-0.079	0.937	
MA-Nonseasonal-01	-0.999978	0.057418	-17.416	< 2e-16	\$~{***}\$
MA-Seasonal-04	0.877926	0.062548	14.036	< 2e-16	\$~{***}\$

Signif. codes: 0 '\$~{***}\$' 0.001 '\$~{**}\$' 0.01 '\$~{*}\$' 0.05 '.' 0.1 ' ' 1

SEATS adj. ARIMA: (3 1 1)(0 1 1) Obs.: 80 Transform: log

AICc: 2474, BIC: 2490 QS (no seasonality in final):0.625

Box-Ljung (no autocorr.): 14.69 Shapiro (normality): 0.9751

APÊNDICE C – DIAGNÓSTICOS DOS TESTES *DICKEY-FULLER* AUMENTADO

Diagnóstico dos testes *Dickey-Fuller* aumentado para as variáveis Receita da Intermediação Financeira (RIF), Depósito Total (DT) e Insumos Exógenos (W) do banco 10045 no intervalo entre o primeiro trimestre de 2000 ao quarto trimestre de 2019.

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression trend (RIF)

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.34187	-0.02519	0.00923	0.03750	0.12098

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.0238897	0.6765941	1.513	0.135
z.lag.1	-0.0609978	0.0401627	-1.519	0.134
tt	0.0009189	0.0007054	1.303	0.197
z.diff.lag	0.0287197	0.0944436	0.304	0.762

Residual standard error: 0.0683 on 65 degrees of freedom

Multiple R-squared: 0.03491, Adjusted R-squared: -0.009632

F-statistic: 0.7838 on 3 and 65 DF, p-value: 0.5073

Value of test-statistic is: -1.5188 0.7776 1.1617

Critical values for test statistics:

1pct 5pct 10pct

```
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression drift (RIF)

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.35605	-0.02622	0.00888	0.04071	0.11845

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.326816	0.416273	0.785	0.435
z.lag.1	-0.018743	0.023810	-0.787	0.434
z.diff.lag	-0.009028	0.090362	-0.100	0.921

Residual standard error: 0.06866 on 66 degrees of freedom

Multiple R-squared: 0.009712, Adjusted R-squared: -0.0203

F-statistic: 0.3236 on 2 and 66 DF, p-value: 0.7247

Value of test-statistic is: -0.7872 0.3144

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.51	-2.89	-2.58
phi1	6.70	4.71	3.86

```
#####
```

```
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

```
Test regression none (RIF)
```

```
Call:
```

```
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
```

```
Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.35397	-0.02739	0.00497	0.04397	0.12098

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	-5.303e-05	4.728e-04	-0.112	0.911
z.diff.lag	-1.484e-02	8.980e-02	-0.165	0.869

```
Residual standard error: 0.06847 on 67 degrees of freedom
```

```
Multiple R-squared: 0.0006413, Adjusted R-squared: -0.02919
```

```
F-statistic: 0.0215 on 2 and 67 DF, p-value: 0.9787
```

```
Value of test-statistic is: -0.1122
```

```
Critical values for test statistics:
```

	1pct	5pct	10pct
tau1	-2.6	-1.95	-1.61

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

```
Test regression trend (DT)
```

```
Call:
```

```
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.108031	-0.032372	-0.003957	0.029713	0.256839

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.9179172	0.6694133	1.371	0.175
z.lag.1	-0.0484898	0.0357477	-1.356	0.180
tt	0.0006325	0.0005791	1.092	0.279
z.diff.lag	0.1437196	0.1228943	1.169	0.246

Residual standard error: 0.05987 on 65 degrees of freedom

Multiple R-squared: 0.04137, Adjusted R-squared: -0.002872

F-statistic: 0.9351 on 3 and 65 DF, p-value: 0.4289

Value of test-statistic is: -1.3564 1.3727 0.9215

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-4.04	-3.45	-3.15
phi2	6.50	4.88	4.16
phi3	8.73	6.49	5.47

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression drift (DT)

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.102328	-0.033211	-0.008653	0.025442	0.273386

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.35893	0.43214	0.831	0.409
z.lag.1	-0.01803	0.02240	-0.805	0.424
z.diff.lag	0.12460	0.12182	1.023	0.310

Residual standard error: 0.05996 on 66 degrees of freedom

Multiple R-squared: 0.02378, Adjusted R-squared: -0.005804

F-statistic: 0.8038 on 2 and 66 DF, p-value: 0.452

Value of test-statistic is: -0.805 1.4583

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.51	-2.89	-2.58
phi1	6.70	4.71	3.86

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression none (DT)

Call:

```
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.104612	-0.033949	-0.004573	0.025387	0.276003

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	0.0005707	0.0003816	1.496	0.139
z.diff.lag	0.1193325	0.1213718	0.983	0.329

Residual standard error: 0.05982 on 67 degrees of freedom
 Multiple R-squared: 0.05623, Adjusted R-squared: 0.02805
 F-statistic: 1.996 on 2 and 67 DF, p-value: 0.1439

Value of test-statistic is: 1.4957

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.6	-1.95	-1.61

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression trend (W)

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.212174	-0.031323	0.002498	0.033107	0.181934

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.1644997	0.8066947	1.444	0.154
z.lag.1	-0.0702030	0.0490880	-1.430	0.157
tt	0.0009626	0.0008813	1.092	0.279
z.diff.lag	0.1756874	0.1177431	1.492	0.141

Residual standard error: 0.06791 on 65 degrees of freedom
 Multiple R-squared: 0.05165, Adjusted R-squared: 0.00788
 F-statistic: 1.18 on 3 and 65 DF, p-value: 0.3242

Value of test-statistic is: -1.4301 0.8196 1.0883

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-4.04	-3.45	-3.15
phi2	6.50	4.88	4.16
phi3	8.73	6.49	5.47

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression drift (W)

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.221973	-0.026063	0.007083	0.036684	0.181703

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.39659	0.39611	1.001	0.320
z.lag.1	-0.02287	0.02310	-0.990	0.326
z.diff.lag	0.13368	0.11145	1.200	0.235

Residual standard error: 0.06801 on 66 degrees of freedom

Multiple R-squared: 0.03424, Adjusted R-squared: 0.004979

F-statistic: 1.17 on 2 and 66 DF, p-value: 0.3167

Value of test-statistic is: -0.9904 0.6311

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.51	-2.89	-2.58
phi1	6.70	4.71	3.86

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression none (W)

Call:

```
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.21463	-0.02910	0.00617	0.03306	0.18426

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	0.0002453	0.0004813	0.510	0.612
z.diff.lag	0.1301213	0.1113934	1.168	0.247

Residual standard error: 0.06801 on 67 degrees of freedom

Multiple R-squared: 0.02624, Adjusted R-squared: -0.002827

F-statistic: 0.9027 on 2 and 67 DF, p-value: 0.4103

Value of test-statistic is: 0.5096

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.6	-1.95	-1.61

APÊNDICE D – TRATAMENTO DOS DADOS APÓS O FILTRO DE SAZONALIDADE E A TRANSFORMAÇÃO LOGARÍTMICA

Código: 20255 – Duas NaN's geradas na variável RIF, com grande intervalo de tempo entre elas. Estratégia de tratamento: O arquivo foi dividido em 4 sub-séries-temporais, isolando os NaN's nos períodos próximos. O primeiro NaN foi substituído pela mediana e o segundo pela média.

Código: 51066 – Três NaN's geradas na variável RIF, duas em sequência e uma com pequeno intervalo de tempo entre elas. Estratégia de tratamento: Duas variáveis foram substituídas pelo log do valor não dessazonalizado e uma substituída pela mediana da amostra total.

Código: 51114 – Duas NaN's geradas na variável RIF e uma na variável ROA. Estratégia de tratamento: Na última linha da amostra estão uma NaN da RIF e uma da ROA. Logo, é eliminada a última linha do código e é feita a substituição da NaN restante pela mediana.

Código: 51183 – Uma NaN na RIF e uma na PC. Estratégia de tratamento: Substituição pela mediana nos dois casos.

Código: 51949 – Gerados ZERO nas variáveis CRD e SOC e Infinito nas variáveis PC e CM. Estratégia de tratamento: Substituição pela mediana nos 4 casos.

Código: 50290345 – Três NaN's na variável RIF, 4 menos infinitos na variável DT, 4, infinitos nas variáveis CRD. Estratégia de tratamento: 2 NaN's e os 4 infinitos estão nas últimas 4 linhas da amostra. Logo, são eliminadas as 4 últimas linhas e a NaN que sobra é substituída pela mediana da variável.

Código: 60498557 – Três NaN's na variável W. O arquivo foi dividido em 4 sub-séries-temporais, isolando os NaN's nos períodos próximos. É observado a inexistência de valores extremos, logo os 3 *outliers* são substituídos pela média do intervalo temporal do subgrupo.

Código: 60518222 – Uma NaN na variável RIF. Estratégia de tratamento: Substituição pela mediana.

Código: 83876003 – Uma NaN na variável ROA. Estratégia de tratamento: Substituição pela mediana.

**APÊNDICE E – ESTIMATIVAS TRIMESTRAIS DO ÍNDICE DE
HERFINDAHL-HIRSCHMANN NORMALIZADO, DO CUSTO MÉDIO, DA
PROXY DO PREÇO DO CRÉDITO E DA ESTATÍSTICA H**

Tabela 15 – Valores estimados do $IHHn$, do $\overline{CM}$, do $\overline{PC}$ e da Estatística H .
entre os anos de 2000 e 2019.

(continua)

	Data	Obs	NAG	TB	RIF	PC	CM	DTT	SOC	ATA	EstHRIF	EstHRT
1	mar/00	115	16359	0	11.991267	0.666963	0.696324	0.120727	0.113807	0.080212		
2	jun/00	116	16393	0	12.252605	0.789832	0.819729	0.116840	0.110472	0.081909	0.932558	0.867166
3	sep/00	112	16060	0	12.042408	0.610681	0.630217	0.117390	0.108364	0.084463	0.766235	0.841235
4	dec/00	110	16530	0	12.278683	0.691936	0.722567	0.116739	0.100310	0.083069	0.777588	0.826962
5	mar/01	107	16725	0	12.168482	0.592003	0.629015	0.120743	0.095366	0.081282	0.322932	0.291601
6	jun/01	107	16913	0	12.368099	0.694746	0.726341	0.119418	0.077599	0.075007	0.798263	0.809269
7	sep/01	105	17183	0	12.545618	0.747230	0.746388	0.117205	0.075118	0.073459	0.672099	0.691136
8	dec/01	104	16975	0	12.454647	0.760206	0.804494	0.113920	0.076347	0.077674	0.643080	0.683393
9	mar/02	100	16988	0	12.317352	0.624607	0.643789	0.118431	0.081318	0.079676	0.569349	0.724148
10	jun/02	100	14314	0	12.673173	0.773448	0.786786	0.113535	0.082291	0.080609	0.947652	0.900428
11	sep/02	100	17128	0	12.842257	0.809743	0.834081	0.117781	0.083572	0.084674	0.678279	0.661112
12	dec/02	97	15170	0	12.766348	0.807208	0.822202	0.117400	0.087716	0.091240	0.360432	0.480887
13	mar/03	97	17303	0	12.448433	0.611068	0.627806	0.115642	0.089939	0.090851	1.001147	0.809841
14	jun/03	97	17275	0	12.324644	0.688884	0.812203	0.120146	0.096163	0.096328	0.203269	0.713220
15	sep/03	96	17375	0	12.374351	0.627681	0.658706	0.119683	0.096676	0.098079	0.883898	0.601943
16	dec/03	94	15414	0	12.393917	0.666940	0.709343	0.121860	0.100191	0.100647	0.526896	0.592206
17	mar/04	91	17235	0	12.283964	0.515990	0.555718	0.120041	0.099730	0.096598	0.621768	0.691400
18	jun/04	90	17396	0	12.409260	0.560693	0.607412	0.117751	0.099128	0.092735	0.811016	0.751090
19	sep/04	90	17402	0	12.162990	0.470585	0.540937	0.115472	0.097859	0.093949	0.403930	0.413162
20	dec/04	90	17446	0	12.208306	0.535133	0.643495	0.115414	0.101292	0.094899	0.860238	0.788157
21	mar/05	90	17641	0	12.353558	0.554235	0.605805	0.115339	0.101677	0.094652	0.641098	0.824045
22	jun/05	90	12971	0	12.359158	0.578915	0.676669	0.112602	0.101414	0.091110	0.622932	0.262185
23	sep/05	90	15683	0	12.378967	0.551641	0.616201	0.112796	0.099843	0.090168	0.416628	0.653762
24	dec/05	89	14745	0	12.488069	0.563659	0.644264	0.114152	0.100587	0.090453	0.358846	0.485805
25	mar/06	89	18021	0	12.474263	0.499722	0.555576	0.113692	0.100377	0.089648	0.895289	0.781244
26	jun/06	88	18162	0	12.530032	0.511447	0.567908	0.110897	0.100884	0.090655	0.671622	0.622720
27	sep/06	87	18283	0	12.461012	0.447667	0.466151	0.113949	0.103195	0.094140	0.548071	0.520225
28	dec/06	87	18411	0	12.531900	0.492505	0.529429	0.115167	0.104985	0.092854	0.892733	0.739430
29	mar/07	87	13365	0	12.516894	0.412272	0.445198	0.115145	0.105689	0.091584	0.131983	0.672430
30	jun/07	85	18611	0	12.612429	0.439475	0.490154	0.113298	0.104263	0.089275	0.903959	0.722436
31	sep/07	84	18718	0	12.532239	0.361836	0.425578	0.113612	0.102624	0.091347	0.755944	0.502845
32	dec/07	84	14999	0	12.648655	0.384851	0.463842	0.115331	0.102731	0.091906	0.679900	0.541672
33	mar/08	82	19101	0	12.701182	0.338140	0.409287	0.114904	0.103576	0.091960	0.779889	0.809026
34	jun/08	83	19193	0	12.637102	0.328563	0.402484	0.110238	0.104461	0.094366	0.575202	0.535594
35	sep/08	83	16328	0	12.689876	0.346434	0.418256	0.450823	0.111284	0.103228	0.321789	0.415955
36	dec/08	82	13755	0	12.822214	0.453879	0.542165	0.529149	0.133259	0.131416	1.155919	0.755549
37	mar/09	79	13963	0	12.449642	0.355005	0.454772	0.158785	0.141066	0.138664	0.268592	0.514606
38	jun/09	80	19610	0	12.312663	0.306489	0.423727	0.154569	0.141461	0.137109	0.309916	0.643783
39	sep/09	81	19998	0	12.127861	0.243295	0.344641	0.151548	0.140826	0.137622	0.362160	0.767858
40	dec/09	80	16022	0	12.319730	0.287842	0.390789	0.150541	0.140461	0.137621	0.609274	0.814364
41	mar/10	80	20526	0	12.435190	0.279703	0.341894	0.150158	0.138485	0.135702	0.248308	0.328653
42	jun/10	80	17200	0	12.504049	0.317499	0.375487	0.273477	0.139962	0.134094	0.900044	0.856819
43	sep/10	80	20138	0	12.289164	0.255845	0.331519	0.559732	0.137626	0.132326	1.264603	0.683529
44	dec/10	79	20217	0	12.448732	0.280056	0.341573	0.064030	0.136803	0.135069	1.063748	0.785793
45	mar/11	79	18065	0	12.416395	0.226439	0.294510	0.145565	0.136465	0.135518	0.994907	0.816143
46	jun/11	78	20450	0	12.527971	0.266886	0.349906	0.144906	0.136989	0.134522	0.697428	0.713865
47	sep/11	78	20816	0	12.894887	0.351713	0.401642	0.146409	0.136835	0.133384	0.643699	0.692336
48	dec/11	78	17006	0	12.837696	0.362822	0.422000	0.149251	0.139384	0.136761	0.770246	0.739898
49	mar/12	78	21798	0	12.604143	0.340754	0.415285	0.148750	0.139494	0.136069	1.182503	0.926425

(conclusão)

	Data	Obs	NAG	TB	RIF	PC	CM	DTT	SOC	ATA	EstHRIF	EstHRT
50	jun/12	76	21885	0	12.737931	0.349084	0.407010	0.152901	0.141509	0.135098	1.349684	0.917628
51	sep/12	76	19758	0	12.469208	0.232497	0.300800	0.155094	0.142823	0.137677	0.712880	0.564131
52	dec/12	75	22560	0	12.515610	0.226784	0.289239	0.157347	0.147367	0.138809	0.880396	0.678871
53	mar/13	76	22784	0	12.355947	0.191671	0.253089	0.158653	0.148187	0.137349	0.293531	0.696945
54	jun/13	76	22924	0	12.534391	0.280921	0.354069	0.158851	0.152255	0.137013	0.402267	0.509407
55	sep/13	73	18185	0	12.465964	0.192683	0.249360	0.160043	0.154734	0.138615	0.680643	0.736726
56	dec/13	72	23284	0	12.569075	0.211571	0.276249	0.160635	0.157319	0.142047	-0.145490	-0.180503
57	mar/14	73	23394	0	12.481013	0.174846	0.231807	0.159925	0.158639	0.141679	0.730812	1.033213
58	jun/14	74	23384	0	12.493001	0.198144	0.259968	0.160582	0.161916	0.144431	0.300777	0.722671
59	sep/14	74	23334	0	12.656330	0.196130	0.240955	0.159337	0.162429	0.143226	0.989761	0.745373
60	dec/14	74	23450	0	12.769140	0.236062	0.294575	0.160052	0.163533	0.141105	-0.021924	-0.103491
61	mar/15	74	23447	1	12.942577	0.257374	0.315537	0.161205	0.163201	0.140031	0.451458	0.488153
62	jun/15	74	15971	1	12.851945	0.255126	0.306854	0.161249	0.165493	0.141094	0.724190	0.867265
63	sep/15	75	20949	1	12.947858	0.265205	0.314226	0.153184	0.159570	0.135293	0.813757	0.670711
64	dec/15	74	23143	1	12.947385	0.311531	0.370197	0.154977	0.163110	0.138637	0.421648	0.657715
65	mar/16	70	19335	2	12.655927	0.230945	0.315249	0.156168	0.166162	0.139911	0.850376	0.770750
66	jun/16	70	22967	2	12.688792	0.221367	0.307329	0.157759	0.169132	0.144060	0.729236	0.920506
67	sep/16	72	22902	3	12.673142	0.212552	0.246043	0.165851	0.172753	0.149288	0.778467	1.034955
68	dec/16	73	22770	3	12.719708	0.240272	0.296006	0.170618	0.175550	0.150538	0.843142	0.802297
69	mar/17	73	17399	3	12.539200	0.188284	0.249045	0.168352	0.175437	0.149688	1.120573	0.839806
70	jun/17	75	21867	3	12.738833	0.231504	0.283424	0.165470	0.175503	0.150878	1.364561	1.038402
71	sep/17	75	18584	3	12.387094	0.161753	0.218233	0.162538	0.176587	0.149900	1.361863	0.951284
72	dec/17	77	21468	4	12.523141	0.233188	0.284879	0.161926	0.175095	0.149840	1.038139	0.662928
73	mar/18	77	18331	5	12.325708	0.182307	0.233536	0.158749	0.173314	0.148605	0.857382	0.993980
74	jun/18	76	21341	4	12.680856	0.299558	0.364391	0.156276	0.171496	0.145770	1.220761	1.332038
75	sep/18	76	17059	4	12.460365	0.211694	0.284507	0.154391	0.169826	0.144017	0.938737	0.859584
76	dec/18	76	21188	4	12.516000	0.242356	0.324613	0.153507	0.168781	0.145438	0.707867	0.751058
77	mar/19	77	18727	6	12.391245	0.262535	0.336506	0.153512	0.166314	0.146184	0.872999	0.826805
78	jun/19	77	21081	7	12.485407	0.237728	0.323206	0.150707	0.164734	0.144730	0.700728	0.870072
79	sep/19	77	20652	7	12.588531	0.213324	0.269807	0.149976	0.162512	0.142707	0.169304	0.337568
80	dec/19	76	20428	7	12.600007	0.245822	0.329063	0.150375	0.161135	0.143426	0.616160	0.473805

Fonte: Resultado da pesquisa.

APÊNDICE F – DIAGNÓSTICO DOS TESTES DE DEPENDÊNCIA TRANSVERSAL, DE *BREUSH-PAGAN* PARA HETEROCEDASTICIDADE E DE CORRELAÇÃO SERIAL EM PAINEL

[1] "Testes de dependência transversal (cross-sectional)"

Pesaran CD test for cross-sectional dependence in panels

data: RIF ~ DT + W + CRD + TB

z = 37.09, p-value < 2.2e-16

alternative hypothesis: cross-sectional dependence

Pesaran CD test for cross-sectional dependence in panels

data: RT ~ DT + W + CRD + TB

z = 39.022, p-value < 2.2e-16

alternative hypothesis: cross-sectional dependence

[1] "Teste de Breush-Pagan para Heterocedasticidade"

studentized Breusch-Pagan test

data: efal_RIF

BP = 10.032, df = 4, p-value = 0.0399

studentized Breusch-Pagan test

data: efal_RT

BP = 26.536, df = 4, p-value = 2.467e-05

[1] "Teste Breush-Godfrey/Wooldridge - Teste de correlação serial em painel"

Breusch-Godfrey/Wooldridge test for serial correlation in panel models

data: RIF ~ DT + W + CRD + TB

chisq = 61.932, df = 1, p-value = 3.555e-15

alternative hypothesis: serial correlation in idiosyncratic errors

Breusch-Godfrey/Wooldridge test for serial correlation in panel models

data: $RT \sim DT + W + CRD + TB$

chisq = 197.62, df = 1, p-value < 2.2e-16

alternative hypothesis: serial correlation in idiosyncratic errors