



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO

INSTITUTO DE CIÊNCIAS ECONÔMICAS

PHELIPE TELES DA SILVA

**Determinantes macroeconômicos do spread bancário brasileiro  
(2011-2017)**

SEROPÉDICA, 2019.

Phelipe Teles da Silva

# **Determinantes macroeconômicos do spread bancário brasileiro (2011-2017)**

Monografia apresentada no curso graduação da Universidade Federal Rural do Rio de Janeiro, Instituto de Ciências Sociais Aplicadas, curso de Economia como requisito parcial para obtenção do título de Bacharel em Economia.

Universidade Federal Rural do Rio de Janeiro

Orientadora: Débora Pimentel

Seropédica

2019

## Sumário

<b>1</b>	<b>Introdução</b>	<b>3</b>
<b>2</b>	<b>Visualização e análise descritiva das variáveis</b>	<b>4</b>
<b>3</b>	<b>Revisão de Literatura</b>	<b>9</b>
3.1	Revisão da literatura teórica	9
3.1.1	O Banco como firma	9
3.1.2	O banco como intermediador financeiro	10
3.2	Revisão da literatura empírica	11
<b>4</b>	<b>Referencial Metodológico</b>	<b>14</b>
4.1	Teste de raiz unitária	14
4.2	Teste de Co-integração	15
4.3	Vetores autorregressivos (VAR)	16
4.4	Função de Impulso-resposta	18
4.5	Modelo vetorial de correção de erros	18
4.6	Teste de Johansen	20
<b>5</b>	<b>Apresentação dos resultados</b>	<b>21</b>
5.1	Testes de raiz unitária	21
5.2	Teste de co-integração uniequacional	22
5.3	Modelo vetorial de correção de erros	23
5.3.1	Seleção da ordem	23
5.3.2	Posto de co-integração	23
5.3.3	Estimação	23
5.3.4	Análise dos resíduos	23
5.3.5	Funções de Impulso-Resposta	23
	<b>Referências</b>	<b>24</b>

# 1 Introdução

O spread bancário tem sido de particular interesse para pesquisadores no Brasil, devido à peculiaridade de ser um dos maiores do mundo, o que pode ser visualizado na Figura 1. Nela, podemos ver os maiores spreads bancários da América Latina, no que se destaca a liderança do Brasil por uma larga margem.

Figura 1 – Top 10 spreads bancários da América Latina em 2017

Fonte: indicador “FR.INR.LNDP” do World Bank. Elaboração própria

Com a bem-sucedida estabilização macroeconômica feita pelo Plano Real, esperava-se que esse nível finalmente começasse a convergir para os padrões internacionais mas, embora ele tenha de fato se reduzido consideravelmente, ainda continuou em um patamar relativamente elevado.

Desde então, uma das principais razões que se atribui a isso é a política monetária, conduzida sob o Regime de Metas de Inflação (RMI) desde 1999, de manter a níveis altos a taxa básica de juros com o objetivo de controlar a inflação, porque ela indexa parte dos títulos da dívida pública brasileira, o que torna a compra de títulos públicos mais atraente que o crédito, uma aplicação naturalmente mais arriscada e menos líquida (MANHIÇA; JORGE, 2012, p. 7).

Mais recentemente, o debate foi reanimado devido à persistente queda da taxa Selic, trazendo consigo a expectativa de queda também do spread. Embora essa queda tenha mesmo se efetivado, seu ritmo foi tido como insatisfatório pelas autoridades (SIMÃO; CAMPOS, 2018). Há um elenco de fatores que podem ajudar a explicar isso, de variáveis micro, como a concorrência, assimetrias de informação sobre os tomadores de crédito e capacidade de recuperação de garantias<sup>1</sup>, a macroeconômicas, como a inadimplência e concentração bancária<sup>2</sup>.

Em um cenário de incerteza macroeconômica, é interessante que se investigue os determinantes macroeconômicos do spread bancário, a fim de, com um melhor entendimento do caso brasileiro, contribuir para a efetividade das políticas de redução do spread, o que se apresenta ainda mais importante na atual situação de estagnação econômica, já que um spread mais baixo poderia ajudar na retomada, ao facilitar o acesso ao crédito dos agentes econômicos e, portanto, a realização de investimentos (MANHIÇA; JORGE, 2012, p. 8) além de, de modo geral, significar um sistema financeiro mais desenvolvido, mais eficiente em seu papel da intermediação financeira.

<sup>1</sup> Ver (BANCO CENTRAL DO BRASIL, 2018, p. 13).

<sup>2</sup> Ver (PINHEIRO; BELLOTTO, 2018)

## 2 Visualização e análise descritiva das variáveis

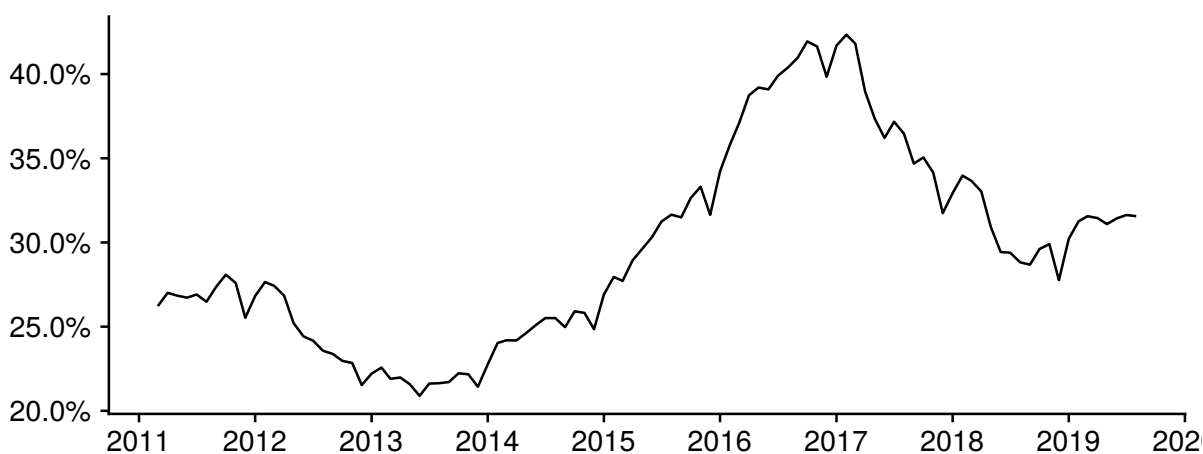
Neste capítulo serão apresentadas as relações teóricas entre a variável dependente e as variáveis independentes incluídas no modelo, assim como a evolução histórica das séries, que cobrem o período de março de 2011 até outubro de 2019.

Começaremos com o spread bancário. Trata-se, mais especificamente, da série 20786 do Sistema Gerenciador de Séries Temporais (SGS) do Banco Central do Brasil, intitulada "Spread médio das operações de crédito com recursos livres - Total", sendo a diferença, em pontos percentuais, entre a taxa média de empréstimo e de captação no mês. Por total, entende-se que ela aglutina operações de pessoas físicas e jurídicas, e por recursos livres, que exclui operações envolvendo taxas regulamentadas, lastreadas em recursos governamentais e afins.

Na literatura, é o que se conhece por spread *ex-ante*, porque é calculado com base não no resultado financeiro efetivo dos bancos, mas nas taxas que eles estabelecem, ou seja, no que eles esperam ganhar. Por isso, essa decisão é consideravelmente influenciada por suas expectativas em relação à economia, sendo portanto mais sensível a mudanças macroeconômicas e ao risco percebido (LEAL, 2007, p. 226). Em contraste, o spread *ex-post* é calculado com base na receita e despesa efetiva advinda da atividade de intermediação financeira (ALMEIDA; DIVINO, 2015, p. 2).

Na Figura 2. 1, podemos ver o aumento do spread que ocorreu a partir de 2014, após uma queda que se iniciou por volta de 2012 e que corresponde às políticas do primeiro governo de Dilma Rousseff com o intuito de reduzir o spread via aumento do portfólio de crédito dos bancos públicos, forçando uma queda pela competição das taxas de juros de empréstimos (ALMEIDA; DIVINO, 2015, p. 1). Com a deterioração das condições macroeconômicas a partir de 2014, o spread entra em escalada, seguida por um declínio iniciado em 2017 por conta da, entre outras coisas, queda persistente da taxa SELIC a partir de então.

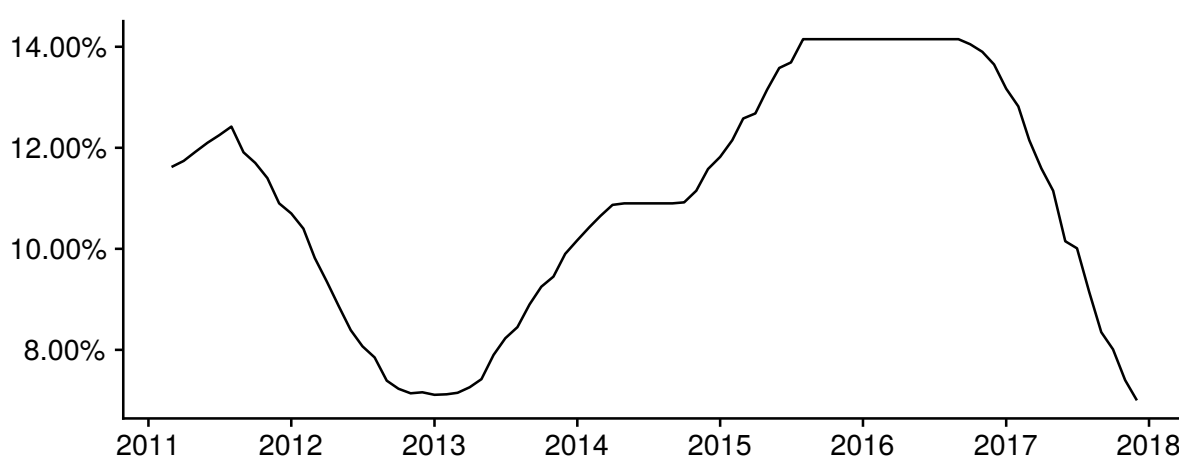
Figura 2 – Spread médio das operações de crédito com recursos livres - Total



Fonte: série 20786 do SGS. Elaboração própria

Este movimento da taxa SELIC pode ser visto na Figura 3. 2. A contraposição das duas séries sugere de imediato que elas são positivamente correlacionadas. De fato, é isso o que naturalmente se argumenta, porque uma taxa básica de juros mais elevada implica maior custo de oportunidade para a atividade de crédito, uma vez que aumenta a rentabilidade dos títulos públicos, tornando mais atrativa uma aplicação que já tem a vantagem de ser mais líquida e menos arriscada (OLIVEIRA; CARVALHO, 2007, p. 372). Por esta razão, é esperado um coeficiente positivo para a taxa básica de juros, pelo efeito custo de oportunidade. Na literatura empírica, não há divergências para essa estimativa, tendo a maior parte dos estudos encontrado a relação esperada (LEAL, 2007, p. 233-234).

Figura 3 – Taxa de juros - Selic acumulada no mês anualizada

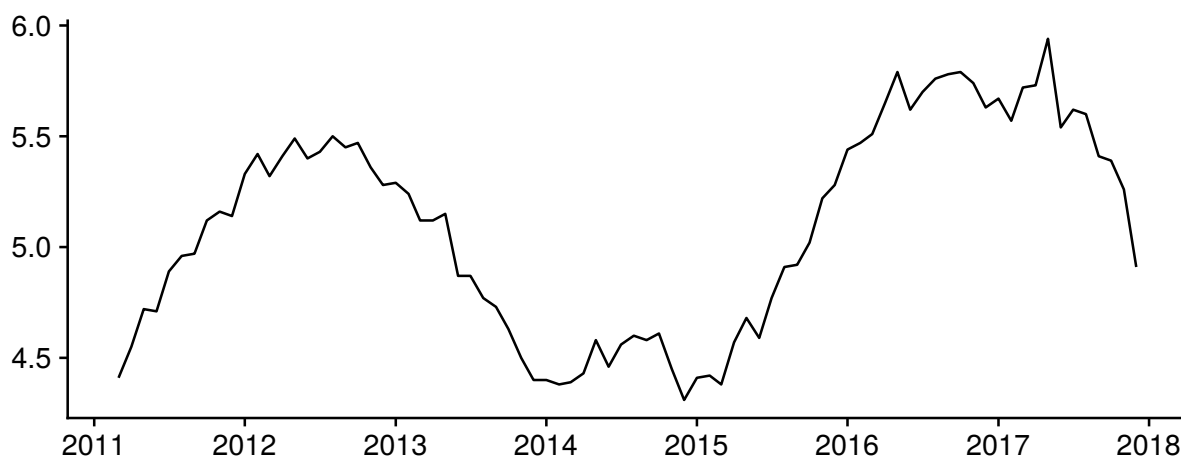


Fonte: série 4189 do SGS. Elaboração própria

Um padrão parecido pode ser observado na série da Inadimplência, na Figura 4, que passou a cair consideravelmente em 2012 para depois aumentar a partir de 2014 e 2015, com o advento da crise econômica. Não é surpreendente que haja aqui uma correlação com a Selic, visto que um aumento da taxa básica de juros eleva o custo de captação, que é então repassado para o juros final, prejudicando a capacidade de pagamento (OLIVEIRA; CARVALHO, 2007, p. 390). A inclusão desta variável, portanto, serve para capturar o efeito do risco de crédito sobre o spread, *ceteris paribus*.

A série de inflação escolhida foi o Índice Geral de Preços (IGP-DI), por ter apresentado resultados mais consistentes que o IPCA (MEDEIROS, 2009, p. 66) e por ser mais abrangente (AFANASIEFF; LHACER; NAKANE, M., 2002, p. 21). A inclusão dessa variável se faz prudente porque sua variação pode influenciar a taxa básica de juros e a política de juros dos bancos, portanto o spread (BIGNOTTO; RODRIGUES, 2006, p. 14). O que se espera obter na estimação é um coeficiente positivo, indicando uma relação direta entre a inflação e o spread, devido ao fato de que em um ambiente econômico não sujeito à instabilidade dos preços os bancos não precisariam se proteger dela via spread. Se por um lado o efeito teórico esperado é claro, o que sairá na estimação é incerto, porque na literatura varia desde insignificante, como em Oreiro

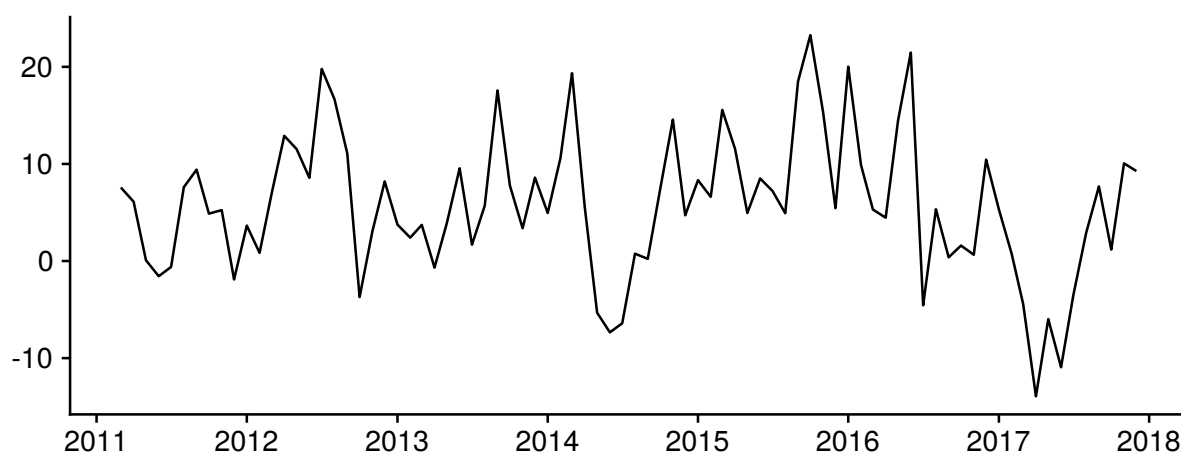
Figura 4 – Inadimplência da carteira de crédito - Total



Fonte: série 21085 do SGS. Elaboração própria

et al. (2006), a um sinal inesperado e significativo, como em Bignotto e Rodrigues (2006) e Afanasieff, Lhacer e M.I Nakane (2002)<sup>3</sup>.

Figura 5 – Índice Geral de Preços (IGP-DI)



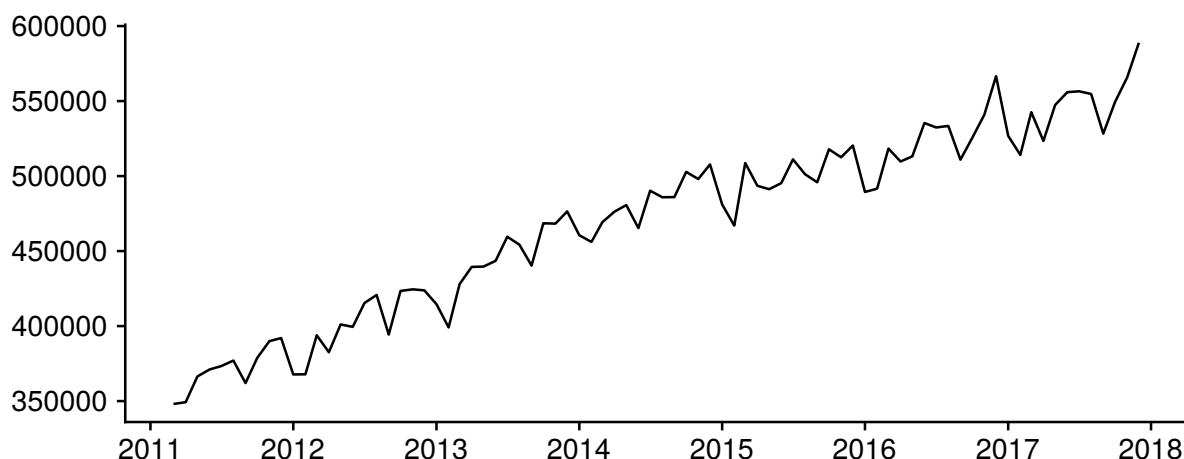
Fonte: IPEA. Elaboração própria

É comum em estudos deste tipo incluir alguma variável para capturar a influência da atividade econômica sobre o spread. Neste trabalho serão considerados dois indicadores. O primeiro é o PIB mensal em valor corrente, obtido pela interpolação dos valores trimestrais divulgados ou projetados. A série está em unidade de milhões e pode ser vista na Figura 6.

O segundo é o Índice de Atividade Econômica do Banco Central (IBC-Br), dessazonalizado e com ano-base em 2002. É útil para ajudar a antecipar o valor mensal do PIB como mensurado pelo IBGE.

<sup>3</sup> Como possível causa, os autores indicam a apropriação de receita de senhoriagem com o spread (AFANASIEFF; LHACER; NAKANE, M., 2002, p. 25)

Figura 6 – PIB Mensal



Fonte: série 4380 do SGS. Elaboração própria

Figura 7 – IBC-Br

Fonte: série *SGS12<sub>1</sub>BCBRDESSAZ12doIPEADATA*. Elaboração própria

Já quanto ao seu efeito sobre o spread, o mais correto é considerá-lo ambíguo. Porque, por um lado, uma maior atividade econômica pode afetar positivamente o spread ao aumentar a taxa de empréstimo via aumento da demanda por crédito, o que dependerá ainda do poder de mercado dos maiores bancos Oreiro et al. (2006, p. 626), e, por outro, pode diminuir o spread ao melhorar as expectativas dos bancos quanto ao risco de crédito Chaim (2013, p. 24). Outro possível efeito é reduzir o spread ao fazer cair a inadimplência, porém isso será controlado com a inclusão da inadimplência no modelo.

Foi ainda considerada a inclusão do Índice de Herfindahl-Hirschmann (IHH) para capturar o efeito da concentração estrutural no Sistema Financeiro Nacional sobre o spread. Como se sabe, este é um indicador do nível de concentração econômica em um mercado, obtido ao somar o quadrado das participações de cada instituição financeira no mercado considerado, que no nosso caso é o mercado de crédito. O BCB considera um IHH entre 0 e 1000 indicativo de baixa concentração, acima de 1000 e menor que 1800, de moderada concentração e acima de 1800 de alta concentração.

No entanto, o índice é divulgado somente trimestralmente e, além disso, termina no ano de 2017 como foi divulgado no Anexo Estatístico do Relatório de Estabilidade Financeira do Banco Central de abril de 2018. Daí se vê que a série apresenta restrições para o tamanho de amostra além de dados faltantes, o que levou à decisão de não incluí-la na estimação, apesar do interesse teórico.

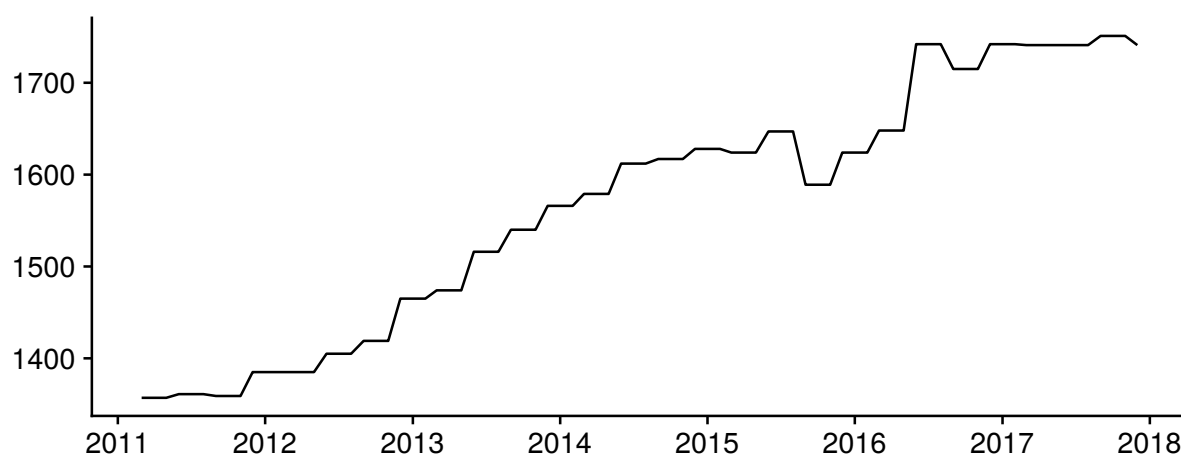
Espera-se, contudo, que isso não signifique incorrer em problemas de viés de variável omitida. Apesar de não ser possível prever ao certo o sinal do coeficiente estimado para esta variável,



acredita-se que o sinal esperado possa ser até mesmo não-significante, com o argumento de que seria a concorrência a variável relevante para explicar o spread, visto que há alta concentração bancária mesmo em países com um sistema financeiro eficiente e bem desenvolvido. Mas também não se descarta a possibilidade do sinal positivo, indicando influência do poder de mercado dos bancos, e até mesmo um sinal negativo, com os investimentos e ganhos de escala associados à concentração aumentando a eficiência da atividade de intermediação, o que puxa o spread para baixo (BANCO CENTRAL DO BRASIL, 2017).

A série do IHH pode ser conferida na Figura 8, do que se conclui que houve um aumento na concentração bancária durante o período. Em contraposição a isso, no Banco Central do Brasil (2017, p. 11) é apontando um aumento da concorrência bancária no mesmo período, o que explicita que nem sempre essas variáveis andam juntas.

Figura 8 – Índice de Herfindahl-Hirschmann - IHH



Fonte: BCB. Elaboração própria

Por fim, a Tabela 1 sumariza o que foi aqui discutido, listando cada variável e seus sinais esperados, além de algumas estatísticas descritivas.

Tabela 1 – Tabela de estatísticas descritivas

Variável	Média	Desvio Padrão	Mín.	Máx.	Sinal esperado
Spread	29,33	6,51	20,89	42,34	
Selic	10,96	2,37	7,00	14,15	+
Inadimplência	5,09	0,47	4,31	5,94	+
IGP-DI	5,74	7,33	-13,93	23,26	+
Atividade Econômica	-2,28	4,86	-13,39	9,70	+/-

Fonte – Elaboração própria.

Nota – Em sinal esperado, “+” indica um coeficiente positivo, “-”, um negativo.

## 3 Revisão de Literatura

### 3.1 Revisão da literatura teórica

Nesta seção serão apresentados dois modelos que se consagraram na literatura sobre a relação entre spread bancário e concentração bancária: o modelo do banco como uma firma maximizadora de lucros de Klein (1971) e o modelo do banco como intermediador financeiro de Ho e Saunders (1981). O intuito é explicitar quais variáveis são relevantes para explicar o spread bancário e por quê, de um ponto de vista macro e microeconômico.

#### 3.1.1 O Banco como firma

Neste modelo, o banco é visto como uma firma que produz serviços voltados para intermediar a oferta e demanda de crédito, receber depósitos ( $D$ ) e fazer empréstimos ( $L$ ), a uma taxa de juros determinada em um mercado monopolístico ou semi-monopolístico, ou seja, em que tem poder de fixar a taxa de juros acima do custo marginal de produção, e é precisamente nesse sentido que o spread bancário é aqui entendido, como o poder de mercado deste banco. (OREIRO et al., 2006)

O autor considera que o banco seja neutro ao risco, buscando maximizar tão somente o valor esperado do lucro, sem considerar sua variância, deparando-se com uma curva de demanda por empréstimos decrescente,  $L(r_L)$ , e uma curva de oferta de depósitos crescente,  $D(r_D)$ , e uma função custo do tipo  $C(D, L)$ , em que  $r_L$  denota a taxa de juros cobrada ao emprestar, e  $r_D$  a paga ao depositante. A função custo, considerando as funções inversas, é assim considerada:

$$\pi(D, L) = r_L(L)L + rM - r_D(D)D - C(D, L) \quad (3.1)$$

em que  $r$  representa a taxa de juros do mercado interbancário, e  $M$ , o que o banco tem disponível para aplicar neste mercado, a saber, tudo o que recebe de depósito e não empresta nem vai para o compulsório (uma taxa  $\alpha$ ),  $M = (1 - \alpha)D - L$ .

E assim, manipulando a equação, obtemos o que representa o resultado da intermediação financeira subtraído de seus custos:

$$\pi(L, D) = (r_L(L) - r)L + (r(1 - \alpha) - r_D(D))D - C(D, L) \quad (3.2)$$

O próximo passo é obter a margem ótima de intermediação, isto é, a que maximiza a função lucro. Para isso, tira-se as derivadas parciais desta função em relação a  $L$  e  $D$  para se chegar às condições de primeira ordem, e, depois de algumas manipulações algébricas, chega-se às equações fundamentais do modelo:

$$\frac{1}{\epsilon_L^*} = \frac{r_L^* - (r - C_L')}{r_L^*} \quad (3.3)$$

$$\frac{1}{\epsilon_D^*} = \frac{r(1 - \alpha) - C_D' - r_D^*}{r_D^*} \quad (3.4)$$

O lado direito da equação é a versão para firma bancária do índice de Lerner<sup>4</sup>, sendo, por definição, em cada um dos casos, igual ao inverso da elasticidade-juros da demanda por empréstimos ( $\epsilon_L^*$ ) e oferta de depósitos ( $\epsilon_D^*$ ). A interpretação desse resultado é que, para maximizar seus lucros, os bancos procuram fixar a taxa de juros num nível acima de seus custos, mas não ao ponto de perderem muitos clientes para a concorrência (o que acontece mais facilmente em mercados com demanda elástica etc.).

Segue-se imediatamente destas equações que o spread bancário será tão maior quanto menos sensíveis forem as elasticidades da demanda por empréstimo e da oferta de depósito em relação à taxa de juros. Outra implicação interessante, e nada óbvia é de que, se a taxa de juros  $r$  do mercado interbancário aumentar, as taxas de intermediação também irão (FREIXAS; ROCHET, 2008, p. 59).

### 3.1.2 O banco como intermediador financeiro

Neste modelo, primeiro apresentado em Ho e Saunders (1981), o banco é considerado como um agente que atua como intermediador entre demandantes e ofertantes de fundos. É um dos mais influentes modelos na literatura, tendo sido estendido por vários outros autores, das quais se falará mais adiante de Maudos e Fernández de Guevara (2004).

Este modelo difere radicalmente do de Klein (1971) no que tange ao tipo de mercado em que o banco atua, não mais harmônico e equilibrado, mas sujeito a incertezas. Neste cenário, o banco não é mais considerado neutro ao risco, mas avesso a ele. Isso porque se depara com duas incertezas: o risco da inadimplência e o risco da taxa de juros, que tem a ver com a possível descoordenação entre demanda por empréstimos e ofertas de depósitos, caso em que o banco terá que recorrer ao mercado interbancário. No caso de demanda excessiva de empréstimos, o banco terá que pedir emprestado e estará sujeito ao risco da taxa de juros aumentar nesse íterim. No caso de oferta excessiva de depósitos, terá que aplicar o excesso, no que estará sujeito ao risco da taxa cair. Maudos e Fernández de Guevara (2004, p. 2262)

Por esta razão, os bancos procuram minimizar esse risco fixando taxas de juros para os depósitos ( $r_D$ ) e para os empréstimos ( $r_L$ ) com a adição de uma pequena margem relativa à taxa de juros do mercado interbancário. Esta margem é o spread ( $s$ ):

<sup>4</sup> Definido como a razão  $(P - C_{mg})/P$ , em que  $P$  é preço e  $C_{mg}$  é o custo marginal, é uma medida do poder de mercado de um agente maximizador de lucros, sendo idêntico à recíproca da elasticidade. (MAUDOS; FERNÁNDEZ DE GUEVARA, 2004)

$$r_D = r - a \quad (3.5)$$

$$r_L = r + b \quad (3.6)$$

$$s = r_L - r_D = a + b \quad (3.7)$$

Para uma derivação das equações do modelo, até a equação do spread ótimo, ver Maudos e Fernández de Guevara (2004, p. 2262) ou Ho e Saunders (1981, p. 584). Para propósitos da revisão, iremos omiti-la, considerando somente suas implicações no que concerne aos determinantes do spread. São eles:

1. Estrutura competitiva do mercado: o que se relaciona com as elasticidades-juros da demanda por empréstimo e da oferta de depósitos, como em Klein (1971).
2. Aversão ao risco: quanto mais avessos, maior o spread.
3. Volatilidade da taxa de juros do mercado interbancário: quanto mais instável a taxa de juros, maior o risco, e portanto maior o spread.
4. Inadimplência: quanto maior o risco de inadimplência, mais o bancos terão que se proteger via spread.
5. Covariância entre risco da taxa de juros e de inadimplência: captura a influência da instabilidade macroeconômica na insolvência das famílias. (OREIRO et al., 2006)
6. Tamanho médio das operações de crédito e de depósito: uma operação de tamanho maior que o médio significa maior perda potencial, da qual se protege com um spread maior.

Esse modelo teórico é a base sob a qual se sustenta a abordagem empírica dos mesmos autores: a estimação do spread em dois estágios. Primeiro, estima-se o spread "puro", usando as variáveis não incluídas no modelo teórico como controle, comumente variáveis microeconômicas e características dos bancos, para então regredi-lo contra as variáveis do modelo teórico Maudos e Fernández de Guevara (2004). Como será visto adiante, uma das vantagens dessa abordagem está em sua flexibilidade, permitindo a inclusão de variáveis sem grandes problemas Almeida e Divino (2015, p. 2).

## 3.2 Revisão da literatura empírica

Esta seção procura revisar os estudos empíricos sobre os determinantes do spread bancário que, cabe antes esclarecer, é calculado com base nas taxas de juros prefixadas cobradas nas modalidades de crédito com recursos livres e na taxa de captação de CDB de trinta dias.

A estabilidade macroeconômica e as políticas governamentais que se seguiram ao Plano Real possibilitaram uma queda significativa do spread bancário, porém somente até um nível

considerado ainda bastante alto para os padrões internacionais. Somando-se a isso a reestruturação bancária observada no período, Afanasieff, Lhacer e M.I Nakane (2002) notam que além da grande variação temporal do spread, há também uma grande variação interbancária. É nesse sentido em que os autores justificam o emprego de técnicas de dados em painel, com o objetivo de capturar essas características do mercado bancário brasileiro para a decomposição dos principais determinantes do spread em variáveis micro e macroeconômicas.

Para isso, os autores aplicam a metodologia de dois passos de Ho e Saunders (1981), usando dados mensais de 142 bancos para o período de fevereiro de 1997 a novembro de 2000, formando um painel com 5578 observações. Para a estimação do spread puro, usam como regressores o intercepto, um vetor temporal e um de características bancárias como controle<sup>5</sup>. Uma vez estimado o spread puro, dado nesse caso pelos coeficientes do vetor temporal e do intercepto, ele é então regredido contra as variáveis macroeconômicas: a taxa de inflação, o compulsório, a taxa SELIC, o crescimento do PIB, a taxa financeira e uma proxy para prêmio de risco.

No primeiro estágio, foi encontrado que o custo operacional, a taxa de depósitos à vista em relação ao ativo total e a de receita de serviços em relação à receita operacional afetam positivamente o spread, e que os bancos estrangeiros estão em média associados a um spread menor. Os autores, ao contraporem o spread puro estimado com o observado, apontam ainda que os fatores microeconômicos não parecem ser os principais determinantes do spread.

Pelos resultados do segundo estágio, há uma relação direta e significativa entre spread e taxa básica de juros, risco, taxa financeira e crescimento do PIB, e contrária no caso da inflação. O efeito estimado do compulsório não foi significativo. Os autores, por fim, concluem que este é um resultado coerente com a o fenômeno de estabilização macroeconômica da época, mas chamam a atenção para a ineficiência e disparidade no mercado bancário, que tornam possível que bancos atuem cobrando taxas muito maiores que seus rivais.

Seguindo linha metodológica parecida, Bignotto e Rodrigues (2006) também se utilizam do modelo teórico de Ho e Saunders (1981) para analisar a influência dos custos de intermediação e dos fatores de risco de crédito e de juros sobre o spread. Para a estimação, fez-se uso de dados de 87 bancos brasileiros ao longo do 1º trimestre de 2001 até o 1º trimestre de 2004, formando uma base de dados em painel desbalanceado com 1131 observações.

Foram usados como regressores o custo administrativo, uma proxy para o risco de crédito e de juros, o market-share, o risco de liquidez, a taxa de receitas de serviços sobre o ativo total e a da despesa tributária sobre o ativo total, o compulsório, a SELIC, a inflação e o ativo total de cada banco corrigido pela inflação, além de uma variável não-observável de aversão ao risco como controle.

O modelo que os autores julgam mais eficiente revela uma relação direta entre o spread

<sup>5</sup> Dentre elas, número de agências, custos operacionais, a taxa de depósitos à vista e a prazo em relação ao ativo total, sendo duas variáveis distintas, a liquidez, pagamento de juros implícitos, alavancagem, uma variável dummy para banco estrangeiro, a receita de serviços sobre a receita operacional total e o patrimônio líquido

e os custos administrativos, o risco de crédito e de juros, o compulsório, a SELIC, e o ativo total, todos com o sinal esperado. Algumas variáveis de controle apresentaram coeficientes inesperados, como no caso da liquidez, da receita de serviços e do market share. Já a variável relativa à carga tributária não se mostrou significativa.

Complementando os importantes estudos feitos pelo Banco Central do Brasil à época, que consistiam na decomposição contábil do spread, Koyama e Márcio I. Nakane (2002) buscaram investigar a sensibilidade do spread a variações nas variáveis que o compõem, isto é, investigar o assunto por uma abordagem econométrica. Para isso, os autores utilizaram a seguinte equação:

$$\ln spread = \beta_0 Tend + \beta_1 \ln selic_t + \beta_2 \ln adm_t + \beta_3 \ln risk_t + \beta_4 \ln imp_t + \beta_5 \ln comp_t$$

Em que *selic* é a taxa básica de juros, *adm* é a razão das despesas administrativas sobre o volume de crédito, *risk* é uma proxy de risco global<sup>6</sup> que busca capturar expectativas, algo mais apropriado quando se trata do spread ex-ante, *imp* é uma medida da incidência de tributos indiretos sobre o spread, *comp* é a taxa de compulsório sobre depósitos à vista exigida pelo Bacen, *Tend* é uma medida de tendência determinista incluída com o intuito de controlar para variáveis que não foram incluídas na equação, como a taxa de inflação, o nível de atividade econômica, e  $\ln$  é o logaritmo natural. Para mais detalhes, ver Koyama e Márcio I. Nakane (2002, p. 10)

Para a estimação da equação de longo prazo, foram usados dados mensais do período de agosto de 1994 a setembro de 2001 em um modelo de vetores autorregressivos (VAR) que, uma vez tratado estatisticamente, revelou que a taxa Selic, as despesas administrativas, o risco e os impostos indiretos afetam positiva e significativamente o spread. Já o compulsório não se mostrou significativo e foi retirado do modelo.

Em seguida, os autores decompuseram o spread do período com base nessa estimação. O resultado indica que todas as variáveis são bem relevantes na composição do spread no período, em especial a variável de risco, relevância esta que foi aumentando com o passar dos anos e que é análoga à considerável participação da inadimplência na decomposição contábil do spread. Os custos administrativos e a Selic vêm em seguida em importância, ambas com tendências sutilmente declinantes. E, por fim, a importância relativa dos impostos indiretos é a menos sobressalente, mas aumentou com os anos.

Um artigo que segue um caminho semelhante em termos de modelagem é o de Oreiro et al. (2006), já que para a estimação dos determinantes também se utilizou um modelo VAR. Este estudo, porém, se distingue por dar foco aos efeitos das variáveis macroeconômicas sobre o spread, cobrindo o período de janeiro de 1995 a dezembro de 2003.

<sup>6</sup> "[É] o spread do rendimento do C-Bond sobre o rendimento do título do Tesouro americano com mesma maturidade." (KOYAMA; NAKANE, M. I., 2002, p. 10)

São usados como regressores séries de frequência mensal como o nível da taxa básica de juros (SELIC) e sua volatilidade (como proxy para o risco de juros), o nível de atividade econômica, o recolhimento compulsório sobre depósitos à vista e a inflação.

Para interpretar os coeficientes estimados, os autores se utilizaram da função de impulso-resposta, em que se procurou observar o efeito de longo prazo que um choque exógeno de um desvio-padrão em uma variável independente tem sobre a variável dependente. A decomposição da variância também foi utilizada, com objetivo similar. Em ambos os casos, busca-se enxergar o efeito estimado do choque ao longo de 12 meses.

Os resultados mostram que um choque exógeno no nível e volatilidade da taxa de juros tem um efeito positivo e persistente sobre o spread, como esperado. Um choque na proxy para o nível de atividade econômica mostrou ter um impacto positivo, o que os autores interpretaram como um possível efeito do poder de mercado dos bancos prevalecendo sobre o efeito inadimplência. O efeito de um choque na inflação se mostrou insignificante estatisticamente. Eles concluem dizendo que um ambiente macroeconômico estável é uma condição indispensável para reduzir o spread.

## 4 Referencial Metodológico

Neste capítulo serão brevemente apresentadas as metodologias utilizadas para a modelagem e tratamento de séries temporais.

Todo estudo do tipo requer primeiro que conheçamos a ordem de integração das séries, isto é, se é estacionária (caso em que é integrada de ordem 0) ou, caso contrário, quantas  $d$  vezes é preciso diferenciá-la para que se torne estacionária (caso em que será integrada de ordem  $d$ ), o que é mais formalmente feito com um teste de raiz unitária.

Uma vez que se saiba quais são as séries não-estacionárias, é prudente investigar se alguma combinação linear delas é estacionária, ou seja, se há co-integração.

Os resultados destes dois testes influenciam enormemente na decisão de como modelar o relacionamento entre as séries.

### 4.1 Teste de raiz unitária

Na análise de séries temporais, é comum nos assegurarmos de que elas sigam um processo estocástico cujas variáveis aleatórias tenham média constante, variância constante e finita e função de autocovariância que dependa somente da defasagem no tempo e não do tempo em si (GUJARATI, 2011, p. 734), isto é, que sejam estacionárias no sentido amplo. Mais formalmente, um processo aleatório formado por variáveis aleatórias discretas  $Y_t$  será estacionário se:

Dado um processo estocástico autorregressivo de ordem 1:

$$Y_t = \rho Y_{t-1} + u_t \quad (4.1)$$

Onde  $u_t \stackrel{iid}{\sim} N(0, \sigma^2)$ . Sabe-se que ele será estacionário se  $|\rho| < 1$ , e não-estacionário se  $|\rho| = 1$ , o que é conhecido como passeio aleatório<sup>7</sup>.

O objetivo do teste de raiz unitária é inferir o valor de  $\rho$  com base nas realizações particulares do processo estocástico. O teste Augmented Dickey-Fuller o faz ao modificar (4.1) como segue:

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_{t-1} - Y_{t-1} + \sum_{i=1}^{p-1} \delta \Delta Y_{t-i} + u_t \\ \Delta Y_t &= \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta \Delta Y_{t-i} + u_t \end{aligned} \quad (4.2)$$

Onde  $\gamma = (\rho - 1)$ . Para logo depois estimar esta equação por Mínimos Quadrados Ordinários (MQO).

O valor estimado de  $\gamma$  serve então para testar a hipótese de raiz unitária  $\gamma = 0$ , que equivale a  $\rho = 1$ , contra a hipótese alternativa unilateral  $\gamma < 0$ , que equivale a  $\rho < 1$ . Calcula-se então o valor t estimado, isto é,  $\hat{\gamma}/EP(\hat{\gamma})$ , com o porém de que sob a hipótese nula  $\gamma = 0$ , essa estatística não segue mais a distribuição t, mas a distribuição  $\tau$ . Isso se dá em razão do teste t ser enviesado no caso de um passeio aleatório (GUJARATI, 2011, p. 748-749).

Importante pontuar que a estatística  $\tau$  é sensível à forma funcional da equação (4.1), isso quer dizer que os valores críticos do teste mudarão se adicionarmos um termo de deslocamento  $\delta$  em (4.1), que denotaremos por  $\tau_\mu$  ou um termo de tendência determinística,  $\beta t$ ,  $\tau_\tau$ .

O termo adicional  $\sum_{i=1}^{p-1} \delta \Delta Y_{t-i}$  em (4.2) serve para generalizar o teste para processos autorregressivos de ordem  $p$ , de forma a preservar a condição de ruído branco em  $u_t$  ao controlar a autocorrelação serial (SOUZA GRÔPPO, 2004, p. 40).

Uma série é dita integrada de ordem 0 quando for diagnosticada como estacionária,  $Y_t \sim I(0)$ , e integrada de ordem  $d$  quando é necessário diferenciá-la  $d$  vezes para torná-la estacionária,  $Y_t \sim I(d)$  (GUJARATI, 2011, p. 740).

## 4.2 Teste de Co-integração

Quando duas ou mais séries são integradas da mesma ordem, mas uma combinação linear delas for integrada de uma ordem menor, dizemos que elas co-integram. O caso comum é quando

<sup>7</sup> Ao resolver recursivamente para o caso quando  $|\rho| = 1$ ,  $Y_t = Y_{t-1} + u_t = Y_0 + \sum u_t$ , pode-se perceber a importante característica da memória infinita, de que ele nunca se esquece do choque, o que também se conhece por tendência estocástica (GUJARATI, 2011, p. 736)



as séries são não-estacionárias, mas uma combinação linear delas é estacionária. Intuitivamente, isso quer dizer que, apesar da tendência estocástica torná-las imprevisíveis individualmente, há um mecanismo de equilíbrio de longo prazo que as impede de se afastarem muito uma da outra (um mecanismo de "correção de erros") (SOUZA GRÔPPO, 2004, p. 46).

Formalmente, as séries  $Y_t$  são co-integradas de ordem  $d, b$  se todas forem integradas de ordem  $d$  e se existe uma combinação linear  $Z_t = \beta' Y_t$  tal que  $Z_t \sim I(d-b), b > 0$  (LÜTKEPOHL, 2005, p. 245), onde  $\beta$  é o vetor cointegrante.

O teste de Engle-Granger procura testar essa condição em duas etapas. Por exemplo, em um contexto bi-variado, em que  $Y_t$  é explicada por  $X_t$ , primeiro estima-se a seguinte equação por MQO:

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

Para na segunda etapa, aplicar o teste ADF nos resíduos desta regressão, ou seja, na equação  $\Delta \hat{u}_t = \gamma \hat{u}_{t-1} + \epsilon_t$ . Se eles forem estacionários, então pode-se concluir que as séries são co-integradas. Vale ressaltar que, como explicado por Gujarati (2011, p. 756), os valores críticos com os quais se deve comparar a estatística de teste calculada são diferentes dos do teste de raiz unitária, porque os resíduos a serem testados foram obtidos com base no  $\beta$  estimado.

Como a relação de co-integração é inerentemente multivariada, este teste não suporta muito bem este caso já que se baseia nos resultados da regressão uniequacional. Além disso, ele depende de como se especifica esta equação, qual variável é endógena e quais são exógenas. Por isso, é apropriado que se investigue a co-integração no contexto multivariado, o que é feito pelo procedimento de Johansen, como veremos (JUSELIUS; HENDRY, 2000, p. 3).

### 4.3 Vetores autorregressivos (VAR)

Segundo Grôppo (2004, p. 51-52), a principal característica de um modelo VAR é o tratamento simétrico das variáveis incluídas no modelo, isto é, não é preciso especificar previamente quais variáveis são exógenas e endógenas, o modelo estrutural que espelha o relacionamento real das variáveis. Por isso, ele está sujeito a poucas restrições teóricas. É somente necessário especificar quais variáveis interagem dentro do sistema e o número de defasagens a serem incluída de forma a melhor capturar esta interação.

São muito úteis para visualizar como uma variável responde a um choque em outra variável, através das funções de impulso-resposta.

Seguindo a notação de Lütkepohl (2005, p. 13), um modelo de vetores autorregressivos de ordem  $p$ , VAR( $p$ ), com  $K$  variáveis, possui a seguinte forma:

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t, \quad t = 0, \pm 1, \pm 2, \dots \quad (4.3)$$

Onde  $y_t = (y_{1t}, \dots, y_{Kt})'$  é um vetor de variáveis aleatórias,  $v = (v_1, \dots, v_K)'$ , um vetor para o intercepto,  $u = (u_1, \dots, u_K)'$ , um vetor de ruído branco gaussiano<sup>8</sup> e  $A_i$  são matrizes  $(K \times K)$  de coeficientes.

Desta equação pode-se perceber que os valores presentes das variáveis são explicados pelas  $p$  defasagens de todas as outras variáveis no sistema.

Dado um AR(1) que começa em  $t = 1$ , resolvendo recursivamente o sistema, chegamos a:

$$\begin{aligned} y_1 &= v + A_1 y_0 + u_1 \\ y_2 &= v + A_2 y_1 + u_2 = v + A_1(v + A_1 y_0 + u_1) + u_2 \\ &= (I_K + A_1)v + A_1^2 y_0 + A_1 u_1 + u_2 \\ y_t &= (I_K + A_1 + \dots + A_1^{t-1})v + A_1^t y_0 + \sum_{i=0}^{t-1} A_1^i u_{t-i} \end{aligned} \quad (4.4)$$

Se todos os autovalores em módulo da matriz  $A_1$  são menores que 1, o que é equivalente a testar se  $\det(I_K - A_1 z) \neq 0$  para  $|z| \leq 1$ , o VAR é dito estável e converge assintoticamente para

$$\begin{aligned} y_t &= (I_K - A_1)^{-1}v + \sum_{i=0}^{\infty} A_1^i u_{t-1-i} \\ y_t &= \mu + \sum_{i=0}^{\infty} A_1^i u_{t-1-i} \end{aligned}$$

O que é também conhecido como a representação em média móvel do modelo VAR. Dela podemos ver que a distribuição de probabilidades de  $y_t$  é unicamente determinada pela distribuição do termo de erro  $u_t$ , que se assume ser gaussiano. Esse resultado pode ser generalizado para AR(p), dado que qualquer processo AR(p) pode ser representado como um processo AR(1) (LÜTKEPOHL, 2005, p. 15). Da equação (4.5) a média e auto-covariâncias de  $y_t$  podem ser derivadas:

$$\begin{aligned} E(y_t) &= \mu, \forall t \\ \Gamma_y(h) &= E(y_t - \mu)(y_{t-h} - \mu)' = \sum_{i=0}^{\infty} A_1^{h+i} \Sigma_u A_1^{i'} \end{aligned}$$

Do que se conclui que um modelo AR(p) estável é estacionário no sentido amplo (LÜTKEPOHL, 2005, p. 15-16)

<sup>8</sup> Logo,  $E(u_t) = 0$  e  $E(u_t u_s') = \Sigma_u$ , onde  $E(u_t u_s') = 0$  para  $t \neq s$  e  $E(u_t u_t') = \Sigma_u, \forall t$ .

## 4.4 Função de Impulso-resposta

A função de impulso-resposta serve para visualizar o efeito provocado por um choque exógeno em uma variável em um sistema dinâmico de equações.

Para vermos o que ocorre matematicamente, seguimos a exposição de Lütkepohl (2005, p. 51-53) e supomos um AR(1) com  $v = 0$ , ou seja,  $y_t = A_1 y_{t-1} + u_t$ :

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = A_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix}$$

Se assumirmos que no período  $t = 0$  há um choque exógeno em  $y_1$ , tal que

$$\begin{bmatrix} y_{1,0} \\ y_{2,0} \\ y_{3,0} \end{bmatrix} = \begin{bmatrix} u_{1,0} \\ u_{2,0} \\ u_{3,0} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Teremos, em  $t$  períodos a frente, que

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = A_1^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Do que se conclui que a resposta do sistema a um impulso na variável  $y_1$  é simplesmente a primeira coluna da matriz  $A_1^t$ , ou, mais geralmente, a resposta a um impulso na variável  $y_k$ , após  $t$  períodos, é a coluna  $k$  da matriz  $A_1^t$ , que, deve-se observar, vai se aproximando de zero quando  $t$  tende ao infinito se o VAR(1) é estável, ou seja, a resposta do sistema vai se dissipando com o tempo.

Quando duas variáveis distintas são correlacionadas no sistema, recomenda-se que se visualize a função de resposta ao impulso ortogonal, porque neste caso um choque não pode ser assumido independente um do outro, sendo provável que afete outra variável no período 0.

## 4.5 Modelo vetorial de correção de erros

Se há alguma relação de equilíbrio entre as variáveis sob análise, gostaríamos de identificá-la, o que normalmente se faz por um teste. Como observado antes, o teste de co-integração uniequacional não é o mais recomendado para um modelo multivariado.

Nesta seção será introduzido o procedimento de Johansen para identificar relações de co-integração em um sistema de equação, mas antes será preciso apresentar o que é o modelo de correção de erros e sua representação em VAR.

Em um modelo de correção de erros, a variação das variáveis depende dos desvios em relação ao equilíbrio entre as variáveis (ao desequilíbrio) e de suas variações defasadas. Por exemplo, dadas duas variáveis  $I(1)$ ,  $y_1$  e  $y_2$ , teremos que:

$$\begin{aligned}\Delta y_{1t} &= \alpha_1(y_{1,t-1} - \beta_1 y_{2,t-1}) + \gamma_{11,1} \Delta y_{1,t-1} + \gamma_{12,1} \Delta y_{2,t-1} + u_{1t} \\ \Delta y_{2t} &= \alpha_2(y_{1,t-1} - \beta_1 y_{2,t-1}) + \gamma_{21,1} \Delta y_{1,t-1} + \gamma_{22,1} \Delta y_{2,t-1} + u_{2t}\end{aligned}\quad (4.5)$$

Por serem  $I(1)$ , as primeiras diferenças  $\Delta y_{i,t}$  são estacionárias e o termo de erro  $u_{it}$  é também assumido estacionário. Colocando o termo de equilíbrio em evidência:

$$\alpha_i(y_{1,t-1} - \beta_1 y_{2,t-1}) = \Delta y_{it} - \gamma_{i1,1} \Delta y_{1,t-1} - \gamma_{i2,1} \Delta y_{2,t-1} - u_{it}$$

Podemos concluir que se algum  $\alpha_{it} \neq 0$ ,  $y_{1t} - \beta y_{2t}$  é estacionária, as duas variáveis co-integram, visto que um processo estacionário não pode ser igual a um processo não-estacionário (LÜTKEPOHL, 2005, p. 244-247).

Em notação matricial, as equações (4.5) tornam-se:

$$\begin{aligned}\Delta y_t &= \alpha \beta' y_{t-1} + \Gamma_1(y_{t-1} - y_{t-2}) + u_t \\ y_t - y_{t-1} &= \alpha \beta' y_{t-1} + \Gamma_1(y_{t-1} - y_{t-2}) + u_t\end{aligned}\quad (4.6)$$

O que pode ser representado por um VAR(2) ao manipular algebricamente (4.6):

$$y_t = (I_K + \Gamma_1 + \alpha \beta') y_{t-1} - \Gamma_1 y_{t-2} + u_t \quad (4.7)$$

Mais geralmente, supondo um processo VAR(p) instável:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t \quad (4.8)$$

O que é verdadeiro se  $\det(I_K - A_1 z - A_2 z^2 - \dots - A_p z^p) = 0$  para  $z = 1$ . Neste caso, conclui-se que a matriz  $\Pi = -(I_K - A_1 - A_2 - \dots - A_p)$  não é inversível e, por isso,  $\text{posto}(\Pi) = r < K$ .

Assim, é possível decompôr a matriz  $\Pi$  em duas matrizes  $\alpha$  e  $\beta$ , de posto  $r$  e dimensão  $(K \times r)$  de maneira que  $\Pi = \alpha \beta'$ .

Podemos manipular a equação (4.8) tirando a primeira diferença (subtraindo  $y_{t-1}$  de ambos os lados) e manipulando algebricamente de modo a obter a matriz  $\Pi = -(I_K - A_1 - \dots - A_p)$  até chegarmos à representação vetorial de correção de erros:

$$\begin{aligned}\Delta y_t &= \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{t-1} \Delta y_{t-p+1} + u_t \\ \Delta y_t &= \alpha \beta' + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{t-1} \Delta y_{t-p+1} + u_t\end{aligned}\quad (4.9)$$

Onde  $\Gamma_i = -(A_{i+1} + \dots + A_p)$ ,  $i = 1, \dots, p-1$ . É possível ainda reescrever (4.9) em sua forma VAR, o que não será aqui demonstrado, ver (LÜTKEPOHL, 2005, p. 248-249).

Agora, ao colocar a matriz  $\Pi$  em evidência na equação (4.9):

$$\begin{aligned}\Pi y_{t-1} &= \Delta y_t - \Gamma_1 \Delta y_{t-1} - \dots - \Gamma_{p-1} \Delta y_{t-p+1} - u_t \\ \alpha \beta' y_{t-1} &= \Delta y_t - \Gamma_1 \Delta y_{t-1} - \dots - \Gamma_{p-1} \Delta y_{t-p+1} - u_t\end{aligned}\quad (4.10)$$

Salta à vista que, se todas as variáveis forem estacionárias após a primeira diferença, ou seja, se são  $I(0)$  ou  $I(1)$ , todos os termos do lado direito da equação são estacionários, logo a matriz  $\Pi y_{t-1} = \alpha \beta' y_{t-1}$  também o é. Isto não muda se retirarmos o termo  $\alpha$  do lado esquerdo ao pre-multiplicar a equação por  $(\alpha' \alpha)^{-1} \alpha'$ , o que nos deixa com  $\beta' y_t$ , que representa uma relação de co-integração entre as variáveis em  $y_t$ .  $\beta$  é conhecida como a matriz de co-integração e  $\alpha$  como a *loading matrix* (LÜTKEPOHL, 2005, p. 244-249).

Quando  $0 < \text{posto}(\Pi) = r < K$ , quer dizer que há  $r$  relações de co-integração no sistema. Já quando  $\text{posto}(\Pi) = 0$ , ou seja, é uma matriz nula,  $\Delta y_t$  é um VAR(p-1) estável, e se  $\text{posto}(\Pi) = K$ , quer dizer que todas as variáveis no sistema são estacionárias e que  $\Delta y_t$  pode ser representado por um VAR(p) estável<sup>9</sup>.

## 4.6 Teste de Johansen

O teste de Johansen, mais detalhadamente explicado em Lütkepohl (2005, p. 327-330), procura testar hipóteses acerca do posto da matriz  $\Pi$  com base numa estatística de teste calculada pela razão entre estimadores de máxima verossimilhança.

O teste tem duas formas. A primeira calcula uma estatística de teste conhecida como estatística de traço e testa as hipóteses:

$$\begin{aligned}H_0 : \text{posto}(\Pi) &= r_0 \\ H_1 : r_0 &< \text{posto}(\Pi) \leq K\end{aligned}\quad (4.11)$$

A segunda é conhecida como estatística do máximo autovalor e testa as hipóteses:

$$\begin{aligned}H_0 : \text{posto}(\Pi) &= r_0 \\ H_1 : \text{posto}(\Pi) &= r_0 + 1\end{aligned}\quad (4.12)$$

Com isso, testam-se as hipóteses nulas de  $\text{posto}(\Pi) = r_0 = 0, 1, \dots, K-1$ , até que não se consiga rejeitar a hipótese nula, do que se conclui que o sistema tem  $r_0$  relações de co-integração.

<sup>9</sup> Isto se justifica porque, neste caso,  $\Pi = (I_K - A_1 - \dots - A_p) = (I_K - A_1 z - \dots - A_p z^p)$  para  $z = 1$ , e, como  $\det(I_K - A_1 - \dots - A_p) \neq 0$ , concluímos que o polinômio característico deste VAR não possui raiz unitária, sendo, portanto, estacionário.

## 5 Apresentação dos resultados

Neste capítulo serão apresentados a etapa de tratamento das séries temporais, os procedimentos estatísticos considerados para a especificação do modelo, os resultados da estimação e, por fim, sua interpretação.

Para explicar o spread bancário, serão incluídas no modelo variáveis para a taxa básica de juros (selic), inadimplência (inad), atividade econômica (pib\_mensal) e inflação (igp), o que pode ser melhor visto na seguinte equação:

$$\text{spread}_t = \beta_0 + \beta_1 \text{selic}_t + \beta_2 \text{inad}_t + \beta_3 \text{pib\_mensal}_t + \beta_4 \text{igp}_t + u_t \quad (5.1)$$

Tratam-se de séries temporais mensais que cobrem o período de março de 2011 a agosto de 2019, somando 102 observações, que podem ser vistas a seguir na Figura 9.

Figura 9 – Séries incluídas no modelo (em log)

Durante todo o capítulo, faz-se uso extensivo do pacote statsmodels do Python, dos autores Seabold e Perktold (2010).

### 5.1 Testes de raiz unitária

Na Tabela 2 estão compilados os resultados do teste de Augmented Dickey-Fuller para a existência de raiz unitária, tendo o número de defasagens sido escolhido com base na minimização do *Bayesian Information Criterion (BIC)*.

Tabela 2 – Teste Augmented Dickey-Fuller de raiz unitária

Variável	Lags	Sem drift e sem tendência		Com drift		Com tendência	
		$\tau$	5%	$\tau_\mu$	5%	$\tau_\tau$	5%
spread	12	-0.05	-1.94	-2.47	-2.89	-4.76*	-3.46
selic	4	-1.15	-1.94	-1.34	-2.89	-1.60	-3.46
pib_mensal	12	2.89	-1.94	-6.09*	-2.89	-3.21	-3.46
igp	0	-4.42*	-1.94	-5.92*	-2.89	-5.89*	-3.45

Fonte – Elaboração própria, com base na função adfuller do pacote statsmodels do Python.

Nota –  $H_0$ : raiz unitária.  $H_a$ : estacionária.

O caractere \* indica evidência de estacionariedade a 5% de significância.

Nela podemos ver que há forte evidência de estacionariedade da série da inflação em qualquer um dos cenários e sem autocorrelação, ou seja, de que se trata de um processo do tipo ruído branco. A variável da atividade econômica foi diagnosticada como estacionária quando modelada como um processo AR(12) com deslocamento. Há, ainda, evidência de estacionariedade para a série do spread quando considerada como um processo AR(12) com tendência.

A Tabela 3 apresenta ainda os resultados para o teste de Phillips-Perron, onde pode-se ver que há evidência de estacionariedade para as variáveis `igp` e `pib_mensal` somente. A função usada lida com a autocorrelação nos resíduos com o estimador de Newey-West. O número de defasagens escolhido é baseado na fórmula  $12 * (n/100)^{(1/4)}$ .

Tabela 3 – Teste Phillips-Perron de raiz unitária

Variável	Lags	Sem drift e sem tendência		Com drift		Com tendência	
		$\tau$	5%	$\tau_\mu$	5%	$\tau_\tau$	5%
<code>spread</code>	13	0.25	-1.94	-1.33	-2.89	-1.61	-3.45
<code>selic</code>	13	-0.95	-1.94	-0.97	-2.89	-1.23	-3.45
<code>pib_mensal</code>	13	3.19	-1.94	-2.18	-2.89	-5.93*	-3.45
<code>igp</code>	13	-4.98*	-1.94	-5.66*	-2.89	-5.62*	-3.45

Fonte – Elaboração própria, com base na função `PhillipsPerron` do pacote `arch` do Python.

Nota –  $H_0$ : raiz unitária.  $H_a$ : estacionária.

O caractere \* indica evidência de estacionariedade a 5% de significância.

Faz-se interessante agora investigar se há co-integração entre alguma combinação linear das variáveis diagnosticadas como não-estacionárias em algum dos cenários.

## 5.2 Teste de co-integração uniequacional

O teste de Engle-Granger foi aplicado sobre a seguinte equação:

$$\text{spread}_t = \beta_0 + \beta_1 t + \beta_2 \text{selic}_t + \beta_3 \text{inad}_t + \beta_4 \text{pib\_mensal}_t + \beta_5 \text{igp}_t + u_t$$

O que se justifica já que é razoável esperar que haja alguma tendência nos resíduos se omitirmos o termo de tendência determinística. Quando aplicado o teste ADF sobre os resíduos dessa regressão, novamente selecionando o número de defasagens, dentre as 12, que minimiza o BIC, não é possível rejeitar a hipótese nula de não co-integração:

Tabela 4 – Teste de co-integração

Engle-Granger	P-valor	1%	5%	10%
-3.61	0.28	-5.19	-4.57	-4.26

Fonte – Elaboração própria, com base na função `coint` do pacote `statsmodels` do Python.

## 5.3 Modelo vetorial de correção de erros

### 5.3.1 Seleção da ordem

Para ajudar nessa decisão, é comum levar em consideração qual ordem minimiza os diversos critérios de seleção disponíveis, como o BIC, o Akaike Information Criterion (AIC), o Hannan-Quinn Information Criterion e o Final Prediction Error (FPE).

Tabela 5 – Seleção da ordem do VECM

	<b>AIC</b>	<b>BIC</b>	<b>FPE</b>	<b>HQIC</b>
<b>0</b>	-29.65	-28.86	1.335e-13	-29.33
<b>1</b>	-30.14	-28.91*	8.166e-14	-29.65
<b>2</b>	-30.54	-28.86	5.550e-14	-29.86
<b>3</b>	-30.81	-28.69	4.267e-14	-29.96
<b>4</b>	-31.12	-28.55	3.210e-14	-30.08
<b>5</b>	-31.30	-28.28	2.764e-14	-30.08*
<b>6</b>	-31.39	-27.92	2.634e-14*	-29.99
<b>7</b>	-31.17	-27.25	3.487e-14	-29.59
<b>8</b>	-31.08	-26.72	4.101e-14	-29.32
<b>9</b>	-31.57	-26.76	2.759e-14	-29.63
<b>10</b>	-31.46	-26.20	3.501e-14	-29.34
<b>11</b>	-31.84	-26.13	2.802e-14	-29.54
<b>12</b>	-32.05*	-25.90	2.761e-14	-29.57

Fonte – Elaboração própria, com base na função `select_order` do pacote `statsmodels` do Python.

Nota – \* destaca o mínimo.

### 5.3.2 Posto de co-integração

### 5.3.3 Estimação

### 5.3.4 Análise dos resíduos

### 5.3.5 Funções de Impulso-Resposta



## Referências

- AFANASIEFF, T.S.; LHACER, P.M.; NAKANE, M.I. The Determinants of Bank Interest Spread in Brazil. **Working Paper Series**, Banco Central do Brasil, v. 46, p. 205–218, Agosto 2002. ISSN 1518-3548.
- ALMEIDA, F.D.; DIVINO, J.A. Determinants of the Banking Spread in the Brazilian Economy: The Role of Micro and Macroeconomic Factors. **International Review of Economics and Finance**, 2015.
- BANCO CENTRAL DO BRASIL. Relatório de Economia Bancária. Brasília, 2017. Disponível em: <[https://www.bcb.gov.br/pec/depep/spread/REB\\_2017.pdf](https://www.bcb.gov.br/pec/depep/spread/REB_2017.pdf)>.
- \_\_\_\_\_. \_\_\_\_\_. Brasília, 2018. Disponível em: <[https://www.bcb.gov.br/content/publicacoes/relatorioeconomiabancaria/reb\\_2018.pdf](https://www.bcb.gov.br/content/publicacoes/relatorioeconomiabancaria/reb_2018.pdf)>.
- BIGNOTTO, Fernando G.; RODRIGUES, Eduardo Augusto de Souza. Fatores de Risco e o Spread bancário no Brasil. **Trabalhos para Discussão**, Banco Central do Brasil, v. 110, p. 1–31, jul. 2006.
- CHAIM, Pedro Luiz Paulino. **Determinantes macroeconômicos do spread bancário brasileiro: uma abordagem autorregressiva**. 2013. Monografia: Faculdade de Economia e Administração – Insper, São Paulo.
- FREIXAS, Xavier; ROCHET, Jean-Charles. **Microeconomics of Banking**. 2. ed. [S.l.]: MIT Press Books, 2008.
- GUJARATI, Damodar. **Econometria Básica**. 5. ed. [S.l.]: Mc Graw Hill, 2011.
- HO, T. S. Y.; SAUNDERS, A. The determinants of bank interest margins: theory and empirical evidence. **Journal of Financial and Quantitative Analysis**, v. 16, p. 581–600, 1981.
- JUSELIUS, Katarina; HENDRY, David. Explaining Cointegration Analysis: Part II. **University of Copenhagen. Department of Economics, Discussion Papers**, v. 22, jan. 2000. DOI: 10.2307/41322908.
- KLEIN, M. A. A theory of banking firm. **Journal of Money, Credit and Banking**, v. 3, p. 205–218, 1971.
- KOYAMA, Sérgio Mikio; NAKANE, Márcio I. Os determinantes do Spread Bancário no Brasil. **Notas Técnicas do Banco Central do Brasil**, Banco Central do Brasil, v. 19, p. 1–14, Abril 2002. ISSN 1519-7212.
- LEAL, Rodrigo Mendes. Estrutura e Determinantes do SPread Bancário no Brasil após 1994: uma ANálise da Literatura Empírica. In: ELSEVIER (Ed.). **Sistema Financeiro: Uma Análise do setor bancário**. Rio de Janeiro: Adriana Amado et. al, 2007. cap. 10, p. 221–254.
- LÜTKEPOHL, Helmut. **New Introduction to Multiple Time Series Analysis**. [S.l.]: Springer Berlin Heidelberg, 2005. DOI: 10.1007/978-3-540-27752-1.

MANHIÇA, Félix Antônio; JORGE, Caroline Teixeira. O nível da taxa básica de juros e o spread bancário no Brasil: uma análise de dados em painel. **Texto para discussão**, IPEA, v. 1710, fev. 2012. ISSN 1415 - 4765. Disponível em: <[http://repositorio.ipea.gov.br/bitstream/11058/1252/1/TD\\_1710.pdf](http://repositorio.ipea.gov.br/bitstream/11058/1252/1/TD_1710.pdf)>.

MAUDOS, Joaquin; FERNÁNDEZ DE GUEVARA, Juan. Factors Explaining the Interest Margin in the Banking Sectors of the European Union. **Journal of Banking & Finance**, v. 28, p. 2259–2281, fev. 2004. DOI: 10.1016/j.jbankfin.2003.09.004.

MEDEIROS, Felipe Rocha. **Um Modelo de determinantes macroeconômicos do spread bancário aplicado aos grandes bancos brasileiros**. [S.l.: s.n.], 2009. Monografia (Bacharel em Ciências Econômicas), Universidade de Brasília (UnB), Brasília, Brasil.

OLIVEIRA, Giuliano Contento; CARVALHO, Carlos Eduardo. O componente custo de oportunidade do spread bancário no Brasil: uma abordagem pós-keynesiana. **Economia e Sociedade**, v. 16, p. 371–404, Dezembro 2007.

OREIRO, José Luís da Costa et al. Determinantes macroeconômicos do spread bancário no Brasil: teoria e evidência recente. pt. **Economia Aplicada**, scielo, v. 10, p. 609–634, dez. 2006. ISSN 1413-8050. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-80502006000400007&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-80502006000400007&nrm=iso)>.

PINHEIRO, Vinicius; BELLOTTO, Alessandra. Concentração dificulta queda mais forte do spread bancário. **Valor Econômico**, abr. 2018. Disponível em: <<https://www.valor.com.br/financas/5440949/concentracao-dificulta-queda-mais-forte-do-spread-bancario>>.

SEABOLD, Skipper; PERKTOLD, Josef. statsmodels: Econometric and statistical modeling with python. In: 9TH Python in Science Conference. [S.l.: s.n.], 2010.

SIMÃO, Edna; CAMPOS, Eduardo. Banco Central quer ver queda mais rápida dos juros bancários. **Valor Econômico**, abr. 2018. Disponível em: <<https://www.valor.com.br/financas/5442475/banco-central-quer-ver-queda-mais-rapida-dos-juros-bancarios>>.

SOUZA GRÔPPO, Gustavo de. **Causalidade das variáveis macroeconômicas sobre o Ibovespa**. 2004. Diss. (Mestrado) – Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo, Estado de São Paulo. DOI: 10.11606/D.11.2004.tde-06012005-165535. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/11/11132/tde-06012005-165535/en.php>>.