

# MENSURANDO E AVALIANDO OS EFEITOS DE UM CHOQUE DE INCERTEZA DA POLÍTICA ECONÔMICA SOBRE A ECONOMIA BRASILEIRA

PIERRE HÍTALO NASCIMENTO SILVA \*  
CÁSSIO DA NÓBREGA BESARRIA †  
MARIA DANIELLA DE OLIVEIRA PEREIRA DA SILVA ‡

## Resumo

Este artigo tem o propósito de criar um índice capaz de mensurar o grau de incerteza da política econômica no Brasil. Esse índice será construído a partir da estimação do sentimento textual contido nas atas de reuniões do Copom, considerando o período de janeiro de 2000 a dezembro de 2018. Posteriormente, analisa-se como um choque de incerteza afeta a dinâmica de um conjunto de variáveis macroeconômicas por meio de um Modelo de Vetores Autorregressivos com restrição de sinais. Foi possível verificar que o aumento da incerteza tem efeitos contracionistas típicos, promovendo redução no consumo e na atividade econômica.

**Palavras-chave:** incerteza econômica; Copom; análise de sentimento; VAR.

## Abstract

This article aims to create an index capable of measuring the degree of uncertainty of economic policy in Brazil. This index will be built from the estimation of the textual sentiment contained in the minutes of Copom meetings, considering the period from January 2000 to December 2018. Subsequently, we analyze how an uncertainty shock affects the dynamics of a set of macroeconomic variables through an Autoregressive Vector Model with signal restriction. It was possible to verify that the increase in uncertainty has contractionary effects, promoting a reduction in consumption and negatively affecting economic activity.

**Keywords:** economic uncertainty; Copom; sentiment analysis; VAR.

**JEL classification:** C83, D8, E32

**DOI:** <http://dx.doi.org/10.11606/1980-5330/ea182617>

---

\* Universidade Federal da Paraíba. Doutorando Programa de Pós-Graduação em Economia. E-mail: pierrehitalo@gmail.com

† Universidade Federal da Paraíba. Professor Programa de Pós-Graduação em Economia. E-mail: cassiodanobrega@yahoo.com.br

‡ Universidade Federal da Paraíba. Professora Departamento de Gestão Pública. E-mail: maria-daniella75@hotmail.com

## 1 Introdução

Os termos risco e incerteza foram tratados como sinônimos pela teoria econômica por muito tempo, Knight (1921) foi o primeiro autor a propor a separação desses conceitos ao apresentar uma definição desagregada. O risco é caracterizado pela capacidade de se atribuir probabilidade, *a priori*, a um determinado evento, mesmo que não seja possível visualizar o resultado futuro<sup>1</sup>, mas que é passível de modelagem. A incerteza é um componente não observável e que se diferencia do risco pelo fato de não ser possível atribuir probabilidade aos eventos futuros, visto que as decisões são tomadas a partir da influência do ambiente para com o agente econômico. Sendo assim, é impossível modelar o conjunto de resultados possíveis/esperados como acontece sob o cenário de risco.

A discussão sobre os fatores geradores da incerteza e suas consequências não é recente, mas novas formas de lidar com esse componente sempre estão em evidência. Algumas pesquisas empíricas têm direcionado o foco para o estudo das consequências trazidas pela incerteza, principalmente em relação ao efeito psicológico dos agentes econômicos. Alchian (1950) argumenta que sob sua presença, julgamentos e opiniões serão divergentes, mesmo embasados nas melhores evidências disponíveis. Em relação ao Brasil, Costa Filho (2014) encontra evidências de que a incerteza apresenta um comportamento anticíclico e de que os períodos prolongados de incerteza são caracterizados por uma atividade econômica desaquecida e, conseqüentemente, uma fase de recessão mais duradoura. A junção desses fatores dificulta a capacidade preditiva dos agentes econômicos e autoridades, fazendo com que esses retardem tomadas de decisões importantes ou as façam de forma incorreta.

Além disso, a presença de incerteza em uma determinada economia traz como consequência o adiamento de muitas das decisões econômicas em função do seu caráter irreversível<sup>2</sup> e do alto grau de volatilidade. Expectativas negativas por parte dos agentes econômicos e empresários acabam sendo formadas e vão sendo empilhadas dentro desse contexto, além de transcenderem de um período para o outro<sup>3</sup> dando origem a um processo retroalimentador. A longo prazo é criada uma memória pessimista que afeta suas escolhas. A expansão duradoura desses choques modifica a estrutura econômica doméstica que fica condicionada a essa nova conjuntura. Além disso, um ambiente tomado por circunstâncias futuras incertas irá refletir um sentimento de indefinição.

Se observado corretamente, o componente incerteza traz informações importantes para a tomada de decisões dos agentes econômicos. Nesse sentido, choques de incerteza acarretam em impactos negativos perante a política monetária, tornando-a mais sensível e/ou comprometendo sua eficiência devido ao seu efeito sobre os agentes econômicos (SCHYMURA, 2017). No que se refere às diretrizes da política econômica, a incapacidade de projeção do cenário econômico de forma clara acarreta na postergação de decisões estratégicas sobre diversos setores.

---

<sup>1</sup> Tome como exemplo o lançamento de um dado, mesmo sem saber qual será o resultado é possível atribuir a probabilidade de ocorrência para cada resultado provável.

<sup>2</sup> Algumas decisões de investimento devido ao tempo de maturidade acabam se tornando irreversíveis.

<sup>3</sup> Considerando a natureza adaptativa das expectativas.

O principal problema reside no fato de a incerteza da política econômica<sup>4</sup> não se tratar de uma variável que possa ser quantificada de maneira usual como ocorre com outras variáveis. Como visto por Knight (1921), ela não apresenta nenhuma distribuição de probabilidade que permita observar e modelar o seu comportamento, fazendo com que, mensurá-la acabe se tornando uma tarefa um pouco árdua. Um dos recursos utilizados para contornar essa dificuldade é a estimação do sentimento de incerteza<sup>5</sup>, uma técnica alternativa que nos últimos anos vem se consolidando na literatura.

Esse método viabiliza a extração da informação ao tentar captá-la mediante a ferramentas de estimação do sentimento textual. Correa *et al.* (2017), construiu um índice capaz de verificar o sentimento de estabilidade financeira a partir dos relatórios de diversos Bancos Centrais. Os autores encontraram indícios de que o sentimento bancário se deteriorava antes de uma crise bancária. Machado *et al.* (2017) encontram evidências de que o tom de incerteza contido nos relatórios de desempenho trimestrais das indústrias influencia a dinâmica do mercado financeiro. Para mais detalhes sobre a discussão de estimação de sentimento ver os trabalhos de: You, DesArmo e Joo (2013), Hutto e Gilbert (2014), Graminho *et al.* (2015), Nopp e Hanbury (2015), Ormerod, Nyman e Tuckett (2015), Silva *et al.* (2017), Machado *et al.* (2017), Aguiar *et al.* (2018) e Li *et al.* (2019).

O sentimento de incerteza, tal como proposto neste artigo, é obtido a partir da análise textual com o auxílio de um dicionário específico que possibilita polarizar os termos. Essa polarização é responsável por exprimir o grau de positividade, negatividade, incerteza, neutralidade, entre outros sentidos ou significados, de uma palavra. Apoiado nessas informações e aplicando métricas que calculam o peso de cada palavra, é possível determinar o sentimento de incerteza da política econômica extraído dos textos selecionados.

Os métodos de *text regression* permitem quantificar palavras e frases de modo a contribuir, por exemplo, com a previsão de variáveis financeiras. Neste ensaio, aplica-se este método com o intuito de estimar o sentimento de incerteza contido nas palavras a partir de bases de dados textuais. Kearney e Liu (2014) enfatizam a existência de muitas fontes de informações prontas para serem “coletadas”: documentos, reportagens, *blogs*, discursos públicos, entre outros. Destaca-se a aplicação do dicionário que permite averiguar o tom contido na escrita concomitantemente com o aprendizado de máquina que promove a conversão de texto em *vector space model* através do uso de matrizes.

O Processamento de Linguagem Natural (NLP) possibilita a interação entre indivíduo e máquina. Ao permitir treiná-la para compreender o vocabulário humano por meio da leitura de textos, o computador é capaz de realizar análise sintática, morfológica, semântica e léxica. O NLP viabiliza a transformação dos dados qualitativos em valores numéricos, por intermédio da estratégia de *bag-of-words* (saco-de-palavras), na qual as palavras passam a ser tratadas como vetores. É possível obter toda a informação textual imprescindível através do *vector space model*, e usá-la como o insumo necessário para analisar o sentimento textual contido nos documentos mediante a transição da palavra para número. Esse processo de extração de informações como sentimento

---

<sup>4</sup>O mesmo acontece com a incerteza econômica.

<sup>5</sup>Em alguns estudos a incerteza é medida pela volatilidade (JURADO; LUDVIGSON; NG, 2015), em outros utilizam volatilidade e estimação do sentimento como em Barboza e Zilberman (2018).

é possível por conta de aprendizagem de máquina baseadas em dicionários.

O sentimento está diretamente atrelado à escrita e, ao selecionar um conjunto de palavras para analisar, consegue-se determinar o seu tom<sup>6</sup>. A partir da tonalidade gerada e mediante o cálculo do peso de cada palavra, é possível obter o parâmetro que expressa o sentimento requerido. Com base nos resultados alcançados, o pesquisador é capaz de fazer uso dessa informação para derivar diferentes indicadores de sentimento: otimismo, pessimismo, felicidade e incerteza, conforme a necessidade do seu problema de pesquisa. Desenvolveu-se, assim, um instrumento de estudo dos padrões de conduta dos agentes econômicos, no qual as respostas obtidas são utilizadas com o intuito de acalmar os ânimos individuais ao fornecer explicações sobre o comportamento da variável analisada, diagnósticos, previsões entre outros.

Contudo, estimar o sentimento “transcrito” ainda se trata de uma metodologia recente e em pleno desenvolvimento. Novas ferramentas, que buscam otimizar cada vez mais o processo, continuam sendo incorporadas. No Brasil, a discussão sobre essa temática ainda é embrionária e se torna ainda mais restrita quando a análise é direcionada para o termo incerteza. É possível destacar alguns trabalhos que fizeram uso dessa técnica no Brasil, sendo eles: Costa Filho (2014), Machado *et al.* (2017), Ferreira *et al.* (2017) e Barboza e Zilberman (2018). Já Godeiro e Oliveira Lima (2017) criaram um Índice de incerteza Macroeconômica agregada para o Brasil a partir da volatilidade, seguindo a metodologia de Jurado, Ludvigson e Ng (2015).

Baseado no arcabouço literário internacional acerca da mensuração da incerteza econômica e seus efeitos sobre a economia, foi desenvolvido por Ferreira *et al.* (2017) o Indicador de Incerteza Econômica – Brasil (IIE-Br)<sup>7</sup>, que apresentou forte relação com os grandes momentos de incerteza vivido pelo país nos últimos anos. Por meio de um estudo econométrico, os autores chegaram à conclusão de que choques de incerteza produzem efeitos negativos sobre a atividade econômica e produção industrial<sup>8</sup>. A construção do IIE-Br é dada pela ponderação dos seguintes componentes<sup>9</sup>: *i*) IIE-Br Mídia: que considera a frequência de notícias à respeito da incerteza nos principais jornais do país<sup>10</sup>; *ii*) IIE-Br Expectativa: as informações dos especialistas referente as expectativas em relação a economia que são divulgadas no Boletim Focus; e *iii*) III-Br Mercado: calculado através da volatilidade do Ibovespa (FERREIRA *et al.*, 2017).

Já a incerteza da política econômica, por sua vez, foi destacada pela primeira vez em Baker, Bloom e Davis (2016)<sup>11</sup>. Os autores criaram um indicador capaz de captar esse sentimento a partir de artigos de jornais, baseado na frequência de palavras associadas a três categorias: incerteza, política e economia. Os autores identificaram uma relação positiva entre incerteza e volatilidade, além de uma relação negativa entre incerteza e o investimento, ou seja, fortes indícios indicam a sua capacidade de afetar variáveis reais da eco-

<sup>6</sup>Neste trabalho, sentimento e tom são tratados como sinônimos.

<sup>7</sup>Divulgado mensalmente pela Fundação Getúlio Vargas.

<sup>8</sup>Ferreira *et al.* (2017) utilizaram as variáveis: IBC-Br e PIM-PF para captar esses efeitos através de um modelo VAR.

<sup>9</sup>Para mais detalhes sobre como ocorre a ponderação consultar Ferreira *et al.* (2017).

<sup>10</sup>Optou-se por utilizar mais de um jornal com o objetivo de corrigir possíveis vies ocasionado por apenas uma mídia.

<sup>11</sup>Os autores são referências nessa temática e atualmente possuem índices de incerteza para diversos países.

nomia. Tomando como base essas hipóteses, Redl (2017) analisou o impacto de choques de incerteza macroeconômica e financeira no Reino Unido a partir de um modelo VAR com restrição de sinais. Os resultados obtidos mostraram que a incerteza política impacta negativamente o PIB e que os efeitos desse choque são superiores aos observados no choque de incerteza financeira.

As atas das reuniões dos Bancos Centrais são os principais instrumentos de comunicação utilizados por esses agentes para apresentar dados macroeconômicos relevantes, tais como: projeções econômicas a respeito da conjuntura doméstica e externa, além de detalhar os procedimentos da tomada de decisão em relação a política monetária. Dado o conteúdo informativo desses documentos, optou-se por trabalhar com as atas do Copom visando o desenvolvimento um índice capaz de mensurar o sentimento de incerteza no Brasil.

As comunicações dos Bancos Centrais surgem como uma fonte promissora de análise do sentimento textual, e sua transparência é vista como fundamental para controlar as expectativas dos agentes econômicos. Montes e Nicolay (2015) realizam um estudo empírico sobre a influência da comunicação do Banco Central do Brasil sobre as expectativas de inflação por meio das atas do Copom. Os autores obtêm indícios de que os especialistas reagem ao “tom” empregado nelas<sup>12</sup> que por sua vez, comprova a eficiência desse instrumento de comunicação utilizado pelo BACEN. Destacam-se alguns autores que realizaram estudos empíricos sobre a comunicação dos Bancos Centrais: Apel e Grimaldi (2012), Bholat *et al.* (2015), Bruno (2016, 2017), Grignani e Fontana (2018) e Rybinski (2018). Na esfera da análise baseada em documentos envolvendo Autoridades Monetárias, Hansen e McMahon (2016) exploram como as informações divulgadas pelo *Federal Open Market Committe* têm efeitos sobre as variáveis reais e de mercado.

Este trabalho é dividido em duas partes complementares. Na primeira etapa serão descritas as fases da construção do índice de incerteza de política econômica, enfatizando o caráter inovador da análise para a discussão sobre sentimento textual no Brasil. A ideia é apresentar um indicador alternativo<sup>13</sup> ao IIE-Br e com menor custo operacional. Posteriormente, optou-se por realizar uma análise comparativa entre o índice desenvolvido e os indicadores de Incerteza Econômica - Brasil (IIE-Br) e o *Economic Policy Uncertainty Index Brazil* (EPU). Teoricamente os índices IIE-Br e IPE mensuram tipos de incerteza diferentes. No entanto, devido à ausência de outros indicadores dentro da temática tratada e aplicados a análise da economia brasileira, utilizou-se esses indicadores para fins comparativos. Além disso, acredita-se que existe uma relação entre ambos, tendo em vista que a incerteza econômica e a incerteza da política econômica estão, em certa medida, interligadas. O EPU está focado em capturar o mesmo tipo de incerteza do IPE, diferenciando apenas na lista de palavras utilizadas e a fonte de coleta dos dados.

Ao comparar o desempenho do IPE com os demais índices, é possível identificar que esses possuem desempenho muito similar, tendo como destaque o compartilhamento de tendências de alta e baixa, correlação positiva, assim como períodos de pico e redução da incerteza. Cabe ressaltar que o IPE conseguiu captar os períodos de maior incerteza doméstica (crises doméstica

<sup>12</sup>Montes e Nicolay (2015) criaram um indicador de percepção da inflação, no qual mostra que os especialistas reagem de forma negativa ao tom empregado do indicador. Se o indicador demonstra que a inflação está sob controle, as expectativas de inflação são reduzidas e *vice-versa*.

<sup>13</sup>A intenção não é substituir, mas sim fornecer aos pesquisadores um indicador capaz de mensurar a incerteza da política econômica, muitas vezes o IIE-Br era utilizado para essa finalidade.

e internacional, períodos eleitorais, entre outros). Posteriormente, como primeiro exercício, observou o impacto de choques de incerteza sobre um conjunto de variáveis macroeconômicas através de um Modelo de Vetores Autorregressivos (VAR) com restrições de sinais, proposto por Uhlig (2005). Nessa etapa de investigação, foi possível constatar que um choque de incerteza de política econômica afeta a economia de forma similar a choque contracionista típico, com redução o consumo e queda na atividade econômica.

## 2 Dados

A primeira fase desse estudo levou em consideração, exclusivamente, as atas do Banco Central. Todas as informações relacionadas às atas do Copom foram reunidas diretamente da base de dados compreendida no portal do BACEN. O critério de seleção levou em conta a disponibilidade desses documentos em inglês e, por isso, optou-se por trabalhar com as informações limitadas ao período janeiro de 2000 a dezembro de 2018.

Deve-se levar em conta uma mudança referente à periodicidade de publicação das atas. Durante o período 2000 a 2005, as reuniões ocorriam de forma mensal, assim como as publicações. A partir do ano de 2006, as reuniões passaram a acontecer a cada quarenta e cinco dias, sendo divulgadas oito atas ao ano. Outro ponto que merece ser destacado é o fato de que em 2002, foram publicadas treze atas ao invés de doze. Devido a essa quebra de periodicidade, passou-se a considerar o período trimestral como base do indicador criado neste artigo, possibilitando que toda a amostra passasse a ter a mesma frequência. As demais séries temporais utilizadas no modelo VAR com restrições de sinais também foram transformadas em dados trimestrais.

A ata do Copom é um dos principais instrumentos de comunicação do Banco Central do Brasil, e apresenta projeções econômicas para o cenário nacional e internacional, controle da inflação e decisões a respeito dos juros. É através dela que a autoridade monetária explica os procedimentos utilizados para a tomada de decisão de política monetária com o objetivo de tornar a comunicação mais transparente e manter sob controle as expectativas (COSTA FILHO; ROCHA, 2010). Desse modo, este estudo se propôs a criar o Índice de Incerteza da Política Econômica (IIPE) derivado a partir das atas do Copom. A série temporal criada nesse estudo considerou o período de 01 de janeiro de 2000 a 31 de dezembro de 2018, totalizando uma amostra de 177 atas que, após transformação de periodicidade em valores trimestrais, resultou em 76 observações.

Em posse desse índice, optou-se por aplicá-lo em um modelo VAR com restrições de sinais com o objetivo de observar a resposta de um conjunto de variáveis macroeconômicas a um choque de incerteza. Os dados utilizados nesta etapa foram coletados através do Sistema Gerenciador de Séries Temporais (SGS) do Banco Central do Brasil, Confederação Nacional da Indústria (CNI), Fundação Getulio Vargas (FGV), Instituto Brasileiro de Geografia e Estatística (IBGE) e Instituto de Pesquisa Econômica Aplicada (IPEA).

### 2.1 Base de dados

A base de dados utilizada neste artigo é composta por observações trimestrais no período que compreende janeiro de 2000 a dezembro de 2018. A taxa Selic é a taxa de juros referencial da economia brasileira e é regulada pelo Copom;

Tabela 1: Descrição dos dados

Variável	Descrição	Fonte	Link
SELIC	Taxa de juros - Meta Selic definida pelo Copom	BACEN	<a href="#">Código 432</a>
IPCA	Índice de preços ao consumidor amplo	IBGE	<a href="#">IPCA</a>
HORAS	Horas trabalhadas na produção industrial	CNI	<a href="#">Horas trabalhadas</a>
CONSUMO	Consumo privado + consumo do governo (em log)	IBGE	<a href="#">Despesas de consumo</a>
FBCF	Formação bruta de capital fixo	IPEA	<a href="#">FBCF</a>
SPREAD	Obtido mediante a diferença entre a Selic e o CDI	BACEN	<a href="#">Códigos 12 e 432</a>
PIB	Produto interno Bruto variação em volume	IBGE	<a href="#">PIB</a>

Fonte:Elaboração própria

o Índice de Preços ao Consumidor Amplo (IPCA) é utilizado como medida de inflação e mede a variação de preços de um conjunto de bens e serviços. A variável horas trabalhadas na indústria é um indicador de desempenho que visa medir a evolução da produção industrial, enquanto a formação bruta de capital fixo (FBCF) mostra o quanto as empresas aumentaram os seus bens de capital, atuando como indicador que mensura o investimento. A variável consumo trata-se da soma das despesas de consumo privada e pública e o produto interno bruto (PIB) considerado foi a série de variação em volume.

Utilizamos como medida de crédito (variável *spread*) a diferença entre a taxa Selic e o CDI. Cabe ressaltar que as variáveis Taxa Preferencial Brasileira (TPB) e Indicador de Custo do Crédito (ICC), refletem com maior precisão o mercado de crédito no Brasil. No entanto, estas não estão disponíveis para o período de análise proposto neste trabalho (1º trimestre de Janeiro de 2000 a 4º trimestre de Dezembro de 2018).

A Tabela 1 mostra as variáveis por nome, descrição, fonte e *link* de acesso.

Destaca-se que a análise empírica da segunda etapa deste artigo está baseada no trabalho proposto por Redl (2017), assim como, a seleção das variáveis.

### 3 Estimação do Sentimento Textual

#### 3.1 IIFE Conjuntura

Definiu-se o  $IIFE_{Conjuntura}$  para abranger a parcela das atas do Copom voltadas para o aspecto conjuntural, em que é observado o cenário econômico do país. Entre as informações divulgadas, destaca-se a decisão da política monetária, voltada para o controle da inflação. A ata é um canal de comunicação direta entre as Autoridades Monetárias e os agentes econômicos, tendo como objetivo tornar mais transparente os fatores que motivaram a decisão do conselho, assim como estabilizar as expectativas dos agentes.

O Processamento de Linguagem Natural permite uma interação entre humano e máquina, e possibilita que sejam determinados os parâmetros que iriam guiar a máquina e permitir a conversão do texto em um *vector space model*<sup>14</sup> que recebeu as *strings* (palavras). A análise das atas do Comitê de Política Monetária demandou a utilização de algumas técnicas específicas de manipulação e mineração textual, de modo a permitir que toda a informação qualitativa necessária contida nos documentos seja extraída e posteriormente transformada em números.

É possível encontrar na literatura algumas metodologias de como estimar o peso das palavras. Por exemplo, Salton e Buckley (1988) utilizam o *vector space model* a partir de uma técnica de normalização dos termos para documentos de diferentes tamanhos visando obter melhores resultados referente ao processo de minimização de *outliers* dentro do vetor corpus<sup>15</sup>.

Neste trabalho, foi escolhida a abordagem proposta por Chisholm e Kolda (1999), onde é aplicada a ponderação nas frequências de palavras, com o objetivo de maximizar a eficiência dos resultados alcançados pelo *vector space model*. O desempenho deste modelo está associado ao esquema de ponderação. Fez-se o uso dos logaritmos com o intuito de ajustar melhor a frequência dos termos em cada ata, impedindo que termos com maiores frequências (*outliers*) venham a ter um peso maior dentro do documento (CHISHOLM; KOLDA, 1999). A estimação do peso das palavras é representada pela equação (1):

$$P_{i,j} \begin{cases} \frac{(1 + \log(Tf_{i,j}))}{(1 + \log(a_j))} \times \log \frac{N}{d_{f,i}} & \text{se } Tf_{i,j} \geq 1 \\ 0 & \text{se } Tf_{i,j} = 0 \end{cases} \quad (1)$$

em que  $P_{i,j}$  representa o peso da palavra  $i$  no documento  $j$ ;  $Tf_i$  é o total de ocorrências de uma palavra  $i$  em um documento  $j$ ;  $a_j$  é a média<sup>16</sup> de frequência das palavras de incerteza presente em um documento  $j$ ;  $N$  é o total de documentos na amostra e;  $d_{f,i}$  representa o total de documentos com ao menos uma ocorrência da palavra  $i$ .

O peso de cada termo pode ser calculado mediante ponderação aplicada a frequência das palavras em cada ata. Pode-se também dividir em duas partes a equação (1). A primeira metade expressa por  $\frac{(1 + \log(Tf_{i,j}))}{(1 + \log(a_j))}$ , representa a parte local da expressão, onde se retrata o peso que a palavra  $i$  tem sobre o documento  $j$ , o termo  $1 + \log(a_j)$  está normalizando<sup>17</sup> a expressão, evitando discrepância entre palavras com frequências muito distintas dentro do mesmo corpo textual. Essa solução de normalização é específica para o *vector space model*. A segunda metade da equação é representada por  $\log \frac{N}{d_{f,i}}$ , e esse termo denota o peso global que a palavra  $i$  tem sobre todo o conjunto de palavras de incerteza.

<sup>14</sup>Representa documentos de texto na forma de vetores de termos, além de permitir modelagem.

<sup>15</sup>O Corpus se trata de uma coletânea restrita de documentos linguísticos manuseados como objeto de estudo que permite modelagem a partir de ferramentas computacionais, para maiores esclarecimentos consultar Aluísio e Almeida (2006).

<sup>16</sup>A média de frequência das palavras de incerteza é obtida mediante a equação:  $a_j = \frac{1}{n} \sum_{i=1}^n FPI_j$ , em que  $n$  representa o número de termos de incerteza propostos por este estudo, que estão presentes no texto;  $FPI$  é a frequência de palavras de incerteza que aparecem no documento  $j$ .

<sup>17</sup>Para maiores esclarecimentos consultar Chisholm e Kolda (1999) os autores apresentam detalhadamente técnicas ponderação para *ovector space model*.



O sentimento textual foi então estimado com base no peso de cada palavra de incerteza contida nas atas de reuniões do Copom. Utilizou-se da abordagem de dicionário através de uma lista de palavras de incerteza previamente determinada. O cálculo do sentimento individual de incerteza da ata é dado por:

$$SI_j = \sum \text{Peso das palavras de incerteza} \quad (2)$$

em que o termo  $SI_j$  é o sentimento textual de incerteza da política econômica da ata  $j$ . A estimação do Índice de Incerteza da Política Econômica se deu mediante o agrupamento desse sentimento individual gerado por cada ata a partir da equação (2).

### 3.2 IPE Projeção

Devido as expectativas do mercado e dos agentes econômicos, se faz necessário que as informações publicadas nas atas das reuniões do Copom sejam transparentes. Divulgar os objetivos da política monetária concomitante com a realização de previsões quantitativas de indicadores relevantes para a tomada de decisão necessita de um grau de transparência, principalmente em relação às perspectivas econômicas e à condução da política monetária no futuro (COSTA FILHO; ROCHA, 2009).

Criou-se o  $IPE_{projecao}$  levando em consideração o aspecto de previsão incluso na ata do Copom. Tornou-se fundamental captar esse sentimento de projeção contido nesses documentos. Ao redigir a ata, as autoridades monetárias deixam claro aquilo que esperam que aconteça tanto no cenário doméstico quanto no internacional, informação essa que afeta diretamente os agentes econômicos.

O processo de criação e estimação do  $IPE_{projecao}$  é idêntico ao do  $IPE_{conjuntura}$  descrito acima, inclusive fazendo uso das Equações (1) e (2), não havendo, portanto, a necessidade de detalhá-lo. A única diferença é em relação a lista de palavras utilizadas no processo de criação<sup>18</sup>: ao trabalhar com o componente de projeção, faz-se necessária a escolha de termos com tempos verbais que transmitem a ideia de projeção e expectativa. Em outras palavras, aquilo que o formulador da política econômica espera que aconteça.

### 3.3 Indicador de Incerteza da Política Econômica

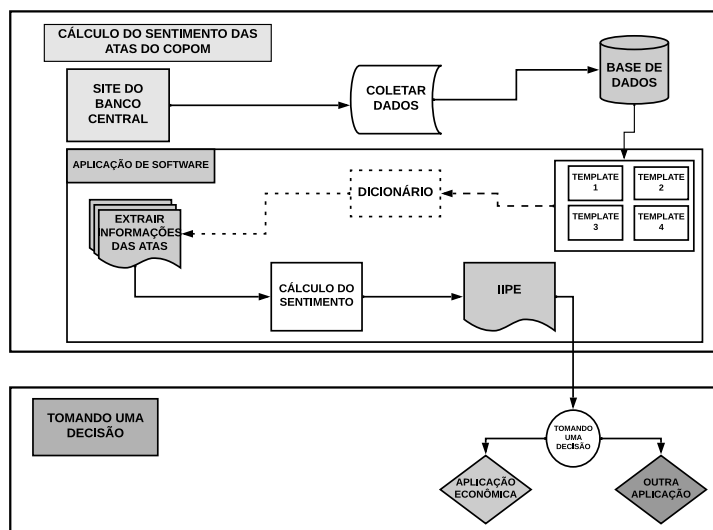
A versão final do Indicador de Incerteza da Política Econômica (IPE) é fruto da combinação do  $IPE_{conjuntura}$  e  $IPE_{projecao}$ , sendo que o peso de cada fator se deu com o objetivo de minimizar possíveis outliers que cada medida venha a ter no indicador final, que é obtido mediante a ponderação:

$$IPE = 0,5IPE_{conjuntura} + 0,5IPE_{projecao} \quad (3)$$

A Figura 7 mostra as diferentes ponderações utilizadas. O objetivo principal é demonstrar que, independente do peso escolhido para o indicador final, o efeito final altera-se minimamente. Ao associar os dois indicadores, obteve-se a série histórica que permite observar o comportamento da incerteza da política econômica no Brasil.

<sup>18</sup>A Figura 1 demonstra claramente como ocorre todo o processo para ambos os índices.

Figura 1: Processo de coleta e estimação do sentimento textual



Fonte: Elaboração própria.

### 3.4 Procedimento de estimação textual

A Figura 1 representa um fluxograma onde se descreve todo o processo automatizado<sup>19</sup> deste artigo. O processo inclui desde a coleta dos dados até o cálculo do sentimento e, posteriormente, a estimação do Índice de Incerteza da Política Econômica, que deriva das atas do Copom. Optou-se por captar manualmente todos os arquivos necessários em PDF diretamente do site do BACEN, criando-se então o banco de dados para trabalho. O objetivo é organizar a base de dados de modo que o acesso aos documentos seja facilitado, além de ficar disponível *off-line* para aplicações e modelagem. Todas as etapas de manuseio e tratamento das atas do Copom foram realizadas a partir do *software R*.

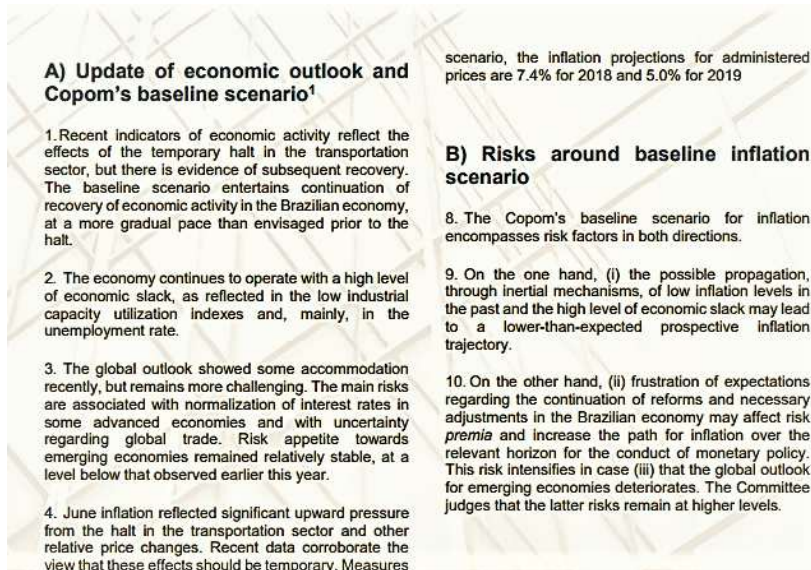
No fluxograma, os *templates* atuam como uma etapa de pré-ajuste, também conhecida como conjunto de regras (remoção de: espaços duplos, pontuação, números, quebra de linha, quebra de página, marca de parágrafo, caracteres maiúsculos), e é aplicada *a priori* sobre a base de dados de modo a prepará-la, a fim de que o algoritmo possa ser executado sem problemas. Uma das maiores dificuldades encontradas no processo de coleta dos dados foi a falta de padronização entre as atas, documentos apresentaram diferença de escrita e de *layout* textual, provocando assim um problema sério, devido ao fato do *software* diferenciar letras maiúsculas e minúsculas, além da organização dos tópicos e exposição da decisão de política monetária. Desta forma, foi possível identificar vários formatos distintos para os documentos que serão analisados ao longo de toda a base textual.

O uso da técnica de *web scraping*<sup>20</sup> é impossibilitado. Como pode ser visto na Figura 2, existe uma quebra textual que impede o algoritmo obter uma enumeração eficiente das palavras de incerteza presente ao longo dos textos. Durante o trabalho com leitura e contagem automatizada de palavras, é pre-

<sup>19</sup>Transformação de informação qualitativa em quantitativa.

<sup>20</sup>Consiste em importar todos os arquivos diretamente do software, dispensando o *download*.

Figura 2: Ata Nº 216: Reunião do Copom - 08 Agosto de 2018



Fonte: Elaboração Banco Central.

ciso que o sistema consiga identificá-las e agrupá-las sem maiores problemas. Por isso, foi preciso tratar toda amostra de modo a criar um padrão entre os documentos.

Terminada esta fase, poucas atas apresentaram problemas ao serem convertidas para arquivos *.txt* (títulos de seções com letras trocadas e o surgimento de símbolos). O algoritmo conseguiu contorná-los pelo fato desses problemas estarem nos títulos e não no corpo textual, e conseqüentemente, não houve interferência na contagem de palavras (mesmo que houvesse, o impacto seria irrisório sobre o resultado). Pao (1978) argumenta que os termos com maior valor semântico se encontram no meio da amostra e não nas extremidades. Todavia, as atas de número 44 e 49 precisaram ser digitadas manualmente devido ao fato de o texto ter ficado codificado<sup>21</sup>, impossibilitando assim, qualquer contagem de frequência.

O dicionário<sup>22</sup> nesse caso não é um termo rígido do algoritmo, podendo ter a sua lista de palavras adaptada ou até mesma substituída conforme necessidade do pesquisador. Finalmente, produz-se como resposta valores numéricos por meio de variáveis definidas e obtidas a partir da extração de informações qualitativas contidas no corpo textual das atas. Decorrida esta etapa, inicia-se o processo de cálculo do sentimento de incerteza de todas as palavras contidas em cada registro coletado. Esse procedimento se repete em *loop*<sup>23</sup> até que o último registro tenha o seu sentimento estimado e, ao final dele, é criado um arquivo único com as informações necessárias. O IPE é então gerado a partir do agrupamento dos sentimentos individuais de cada ata, e ao fim de todo o processo é fornecido um arquivo estruturado no formato *.CSV* com os dados

<sup>21</sup>Surgimento de símbolos ao invés do texto esperado o que impedia o *software*. de conseguir lê-las.

<sup>22</sup>Neste trabalho, dicionário e lista de palavras são tratados como sinônimos

<sup>23</sup>O algoritmo ocorre de forma única para cada documento, criando um laço até que todos passem por ele.

do índice.

### 3.5 Dicionário

#### Procedimento de Organização da Base de Dados e Criação da Lista de Palavras de Incerteza

Dentro das escolhas realizadas em *text regression*, o dicionário é necessário para extrair o sentimento de incerteza dos documentos. Optou-se por trabalhar com as atas em inglês por dois motivos: *i*) o dicionário mais utilizado na análise de sentimento encontrar-se nesse idioma<sup>24</sup>; *ii*) ainda é muito escasso os dicionários escritos em português, atualmente, dentro da proposta deste trabalho existe apenas o dicionário utilizado por Baker, Bloom e Davis (2016) para criar o EPU Brasil a partir das matérias do Jornal Folha de São Paulo. Ao fazer a aplicação do dicionário, permitiu-se utilizar a técnica de *clustering*<sup>25</sup>, que consiste em agrupar um conjunto de objeto, nesse caso, palavras.

O dicionário Harvard-IV<sup>26</sup> (*Harvard Psychosocial Dictionary*), empregado anteriormente como referência para esse tipo de análise, classificava erroneamente o tom de alguns termos financeiros. Devido a este problema, Loughran e McDonald (2011) desenvolveram um dicionário específico capaz de classificar melhor o tom das palavras contidas em textos financeiros. Os autores mencionados foram pioneiros ao aplicar uma nova metodologia de análise textual.

A hipótese proposta por este artigo demandou a criação de uma lista de palavras próprias que englobasse o contexto de incerteza da política econômica no Brasil. A partir desta lista e concomitantemente com a aplicação de um algoritmo sobre as atas do Comitê de Política de Política Monetária, pôde-se construir o Índice de Incerteza da Política Econômica para o Brasil.

Desenvolveu-se, especificamente para o Brasil, um dicionário derivado de uma fusão entre termos retirados da lista de palavras de incerteza definidas pelo Loughran e McDonald (2011), combinada com algumas sentenças escolhidas por Baker, Bloom e Davis (2016), que as subdividiu em três categorias em seu estudo para criação do seu Índice de Incerteza da Política Econômica, além de acrescentar alguns termos próprios relevantes para o contexto brasileiro.

A escolha desses termos se deu mediante a leitura de diversas atas. Percebeu-se a existência de um grupo de palavras significativos, quer seja por sua presença sazonal ligada a acontecimentos políticos expressivos para o país, ou por simplesmente atuarem como bigramas fazendo referência à incerteza. A inclusão dessas palavras “exclusivas” se fez necessária no cenário brasileiro, devido a toda conjuntura peculiar que tem pairado sobre o país nos últimos anos. A adição desses termos foi feita unicamente com o propósito de garantir

<sup>24</sup>Devido à falta de tempo hábil não foi possível desenvolver um dicionário em português capaz de atender à necessidade desta pesquisa. O Processo de construção de um dicionário de forma integral demanda bastante tempo e recurso, é preciso passar por um rígido processo de validação composto por várias etapas. Não existe um consenso na literatura à respeito de qual seria o melhor processo, o que há são trabalhos que apontam para a sabedoria das multidões concomitante com algumas regras adicionais (LOUGHRAN; MCDONALD, 2011; BAKER; BLOOM; DAVIS, 2016; HASELMAYER; JENNY, 2017; CORREA *et al.*, 2017). Tendo em vista a limitação enfrentada, optou-se apenas por validar os seis termos necessários a construção do IIPE. O procedimento de validação desses termos pode ser conferido na subseção 3.5.

<sup>25</sup>Técnica que permite agrupar e contar palavras.

<sup>26</sup>Dicionário de Psicologia.

uma melhor discussão acerca do papel que a incerteza desempenha no comportamento das variáveis macroeconômicas. As palavras próprias que foram incorporadas na lista de incerteza para o Brasil foram: *readjustment, readjustments, reform, reforms, shock e shocks*.

A estruturação da lista de palavras por este trabalho deriva dos dicionários citados anteriormente, além de seguir a divisão de categorias adotada por Baker, Bloom e Davis (2016). A decisão de filtrar as sentenças adotadas neste estudo se deu após a leitura das respectivas listas de palavras, das atas do Copom e da aplicação de alguns testes que serão discutidos na próxima sessão. Chegou-se à conclusão que muitas dessas sentenças eram irrelevantes, e por isso optou-se por excluí-las. O dicionário elegido para a construção do Índice Incerteza da Política Econômica Conjuntura é composto por 45 termos subdivididos em três categorias: incerteza que corresponde a 59% do total de palavras, política que representa 21% e econômica com 20%.

Dos termos presentes no IPE, 68,89% corresponde a sentenças retiradas do dicionário do Loughran e McDonald (2011) que foram totalmente alocadas na categoria incerteza devido a sua lista ser totalmente composta por palavras de incerteza. 13,33%<sup>27</sup> e 4,44% correspondem as palavras utilizadas pelo Baker, Bloom e Davis (2016) para criar os índices de incerteza da política econômica para o Brasil e Estados Unidos respectivamente, que foram destinadas as categorias política e econômica. Por fim; 13,33% das palavras são termos próprios. Após uma análise da conjuntura econômica vivida pelo país, da frequência delas ao longo da série temporal e o teor textual de algumas atas que remetiam a quem lê-se o sentido de incerteza, esses termos foram encaixados nas categorias política e econômica.

Por sua vez, o *IPE<sub>projecao</sub>* é composto por 43 termos retirados unicamente do dicionário do Loughran e McDonald (2011), de início foi realizada a leitura de uma amostra das atas e, a partir dessa leitura, foi pré-selecionado um conjunto de termos que remetiam a projeção. Posteriormente, buscou-se esses termos na relação das palavras de incerteza do dicionário L&M, selecionando as palavras presentes em ambas.

#### Procedimento de Validação da Lista de Palavras

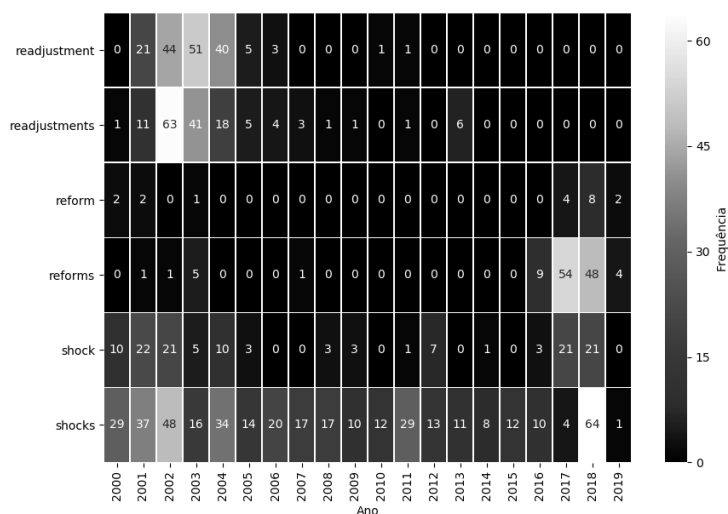
Correa *et al.* (2017) ao estudar a relação entre ciclo financeiro e o sentimento contido nos relatórios de estabilidade financeira, elaborou uma lista de palavras específica para o contexto com 391 palavras<sup>28</sup>, das quais mais de 30% dos termos não constavam no Loughran e McDonald, o que corresponde a um total de pelo menos 117 palavras criadas exclusivamente para o trabalho. Como forma de validar essas sentenças que não integravam o dicionário do Loughran e McDonald<sup>29</sup>, foram escolhidas 25 frases dos relatórios de estabilidade financeira que incluíam esses termos para que dois leitores independentes as classificassem de acordo com sua conotação. Caso houvesse discordância entre eles, uma equipe adicional de mais dois leitores entraria na discussão.

<sup>27</sup>A respeito de palavras repetidas considerou apenas a primeira contagem na seguinte ordem de aparecimento: Loughran e McDonald, Barker lista Brasil e Barker lista USA, evitando contar duas vezes a mesma palavra.

<sup>28</sup>96 positivas e 295 negativas.

<sup>29</sup>São os autores referência na temática de análise de sentimento textual através do uso de dicionário, para maiores esclarecimentos consultar Loughran e McDonald (2011).

Figura 3: Mapa de Calor Frequência dos Termos Próprios



Fonte: Elaboração própria.

O processo de validação dos termos próprios aconteceu em três etapas: *a*) criou-se um mapa de calor apenas com essas sentenças a fim de acompanhar sua evolução ano a ano, *b*) a partir dos dados obtidos em “a” utilizou-se um algoritmo que capta a frequência de buscar realizadas no Google no país durante períodos específicos, e *c*) consistiu em gerar o IIPE utilizando as listas de palavras de incerteza utilizadas pelo Loughran e McDonald (2011) e Baker, Bloom e Davis (2016). Os resultados obtidos ao longo do estágio de validação permitiram obter informações que comprovam que a inclusão desses termos não foi uma ação aleatória, mas sim um processo pensado e previamente definido com objetivo de aumentar o poder do indicador em captar com precisão o sentimento de incerteza.

A Figura 3 mostra o mapa de calor (*heatmap*) dessas palavras nas atas das reuniões ao longo do período observado. Optou-se por usar essa técnica devido a sua clareza. Wilkinson e Friendly (2009) argumentam que, ao utilizar essa ferramenta, é possível compactar matrizes com uma grande quantidade de dados<sup>30</sup> podem ser exibidas de forma compacta facilitando sua inspeção.

Ao analisar a *heatmap* percebemos a presença de palavras sazonais. Nos primeiros anos da amostra, os termos *readjustment* e *readjustments* apareceram de forma mais incisiva nos textos das atas, muito devido à transição de governo e à forma diferente como cada um conduzia a política econômica. Posteriormente, ao observar a parte final, notou-se que as sentenças *reform* e *reforms* apresentam-se de modo significativo; após o *impeachment*. O atual contexto do Brasil favoreceu sua inclusão ao considerá-las vocábulos que se aplicam ao grupo política do IIPE, tendo em vista que o Governo defende a realização de algumas reformas a respeito de questões<sup>31</sup> que trazem impactos significativos para a economia. A indefinição a respeito de sua ocorrência, por sua vez, sinaliza uma mensagem de incerteza para os especialistas. Por fim, as palavras *shock* e *shocks* presentes ao longo de toda série temporal devem ser

<sup>30</sup>Número muito grande de colunas e linhas.

<sup>31</sup>Reforma da previdência e reforma trabalhista.

consideradas como bigramas<sup>32</sup>, ou seja: sozinhas elas possuem um efeito neutro, mas quando precedem outro termo como por exemplo, “choque de juros”, trazem alusão de incerteza. É preciso levar em conta o fato de que para estudar o ambiente brasileiro, é necessário utilizar alguns vocábulos próprios que permitem examinar de forma eficiente o esse sentimento.

A segunda etapa de validação consistiu em aplicar as informações obtidas através do mapa de calor. Mediante uso de um algoritmo foi possível verificar a relevância de alguns termos pertinentes para o nosso estudo pesquisados no Google. No [Apêndice B](#), são apresentados alguns gráficos que refletem o interesse ao longo do tempo no Brasil de alguns bigramas derivados a partir de uma combinação entre as informações das atas e da [Figura 3](#). Mesmo considerando o fato de que os leitores podem ter tido conhecimento destas sentenças a partir de outras fontes e não apenas das atas, é possível notar claramente que o comportamento dos dados reflete o que foi apresentado no mapa de calor<sup>33</sup>.

Como pôde ser visto no fluxograma, o dicionário não é um termo rígido, podendo ser facilmente substituído para a estimação do índice. Com a finalidade de validar a lista de palavras utilizadas por esse estudo, estimou-se novamente o IPE três vezes com o dicionário [Loughran e McDonald \(2011\)](#) e com a lista de palavras utilizadas pelo [Baker, Bloom e Davis \(2016\)](#) para o Brasil e Estados Unidos. Após a observação dos dados, percebeu-se que o comportamento do índice praticamente não sofre alterações, comprovando a robustez da lista de palavras utilizada por esse estudo<sup>34</sup>.

À respeito do  $IPE_{Projecao}$ , nenhum procedimento de validação da lista de palavras foi necessário, tendo em vista que todos os 43 termos foram retirados do dicionário do [Loughran e McDonald \(2011\)](#), o que por si só garante a sua validação.

## 4 Modelo Empírico

### 4.1 VAR

Modelo de Vetores Autorregressivo em sua forma estrutural pode ser representado da seguinte forma:

$$A_0 y_t = \sum_{\ell=1}^p A_{\ell} y_{t-\ell} + \epsilon_t \text{ para } t = 1, \dots, T \quad (4)$$

em que:  $y_t$  é um vetor coluna  $n$ -dimensional de variáveis endógenas,  $A_0$  é uma matriz  $n \times n$  de impactos contemporâneos,  $A_{\ell}$  são  $n \times n$  matrizes de parâmetro de variáveis defasadas,  $\epsilon_t$  é um processo de ruído branco,  $\epsilon_t \sim \mathcal{N}(0, \Sigma_{\epsilon})$ ,  $p$  é o número de atrasos e  $T$  o tamanho da amostra.

Ao multiplicar a Equação (4) por  $A_0^{-1}$  obtém-se a forma reduzida usual do VAR:

$$y_t = \sum_{\ell=1}^p B_{\ell} y_{t-\ell} + u_t \text{ para } t = 1, \dots, T \quad (5)$$

<sup>32</sup>Pares ordenados de palavras.

<sup>33</sup>Para maiores esclarecimentos à respeito do processo de justificação e validação dos termos próprios contatar um dos autores.

<sup>34</sup>Os resultados desse teste de robustez não estão sendo representados nesse trabalho, para qualquer esclarecimento a respeito deles contatar um dos autores.

onde:  $B_\ell = A_0^{-1}A_\ell$  para  $\ell = 1, 2, \dots, p$ ;  $u_t = A_0^{-1}\epsilon_t$  e  $\Omega = E[u_t u_t']$  é a matriz de variância-covariância dos resíduos.

O sistema obtido em (5) é o que deve ser estimado. A partir dos erros de forma reduzida,  $u_t$  é possível conseguir as funções de respostas aos impulsos ortogonais.

#### VAR com restrição de sinais

Um dos benefícios de fazer uso do método de restrição de sinais é o fato de que não é necessário impor restrições sobre toda a matriz  $A_0$ . O propósito de aplicar a abordagem da restrição de sinais é tornar possível a identificação dos choques estruturais, ao impor sobre algumas variáveis uma restrição ao sinal da resposta impulsiva, ao passo que, para as variáveis de interesse, é possível manter o sinal irrestrito com o objetivo de observar a sua resposta ao impulso. Os choques estruturais identificados foram: choque de produtividade, choque na política monetária e choque de incerteza.

Entendendo melhor o funcionamento dessa estratégia de identificação, denote  $a_k$  o vetor de coluna  $k$  da matriz  $A$ , de modo que  $AA' = \Sigma$  e  $A_k$  o vetor de impulso para a  $k$ -ésima variável. O objetivo é identificar apenas  $m$  choques em que  $m \leq n$ . Essa matriz de impulso  $[a^{(1)}, \dots, a^{(m)}]$  pode ser representada como sendo o produto:

$$[a^{(1)}, \dots, a^{(m)}]_{(m \times n)} = \tilde{A}\Lambda \quad (6)$$

em que:  $\Lambda = [\lambda^{(1)}, \dots, \lambda^{(m)}]$  é uma matriz ortonormal  $m \times n$  de modo que:  $\Lambda\Lambda' = I_m$  e  $\tilde{A}$  é um fator Cholesky triangular inferior de  $\Sigma$ .

Seja  $a = a^{(s)}$ ,  $s = 1, 2, \dots, m$  uma das colunas de impulso da matriz  $\tilde{A}\Lambda$  e  $\lambda = \lambda^{(s)} = \tilde{A}^{-1}a$  seja a coluna correspondente de  $\Lambda$ . Uhlig (2005) demonstra que qualquer resposta de impulso  $a$ , pode ser obtida mediante uma combinação linear das respostas de impulso perante uma decomposição de Cholesky  $\Sigma$  que satisfaz  $\Sigma = E[u_t u_t'] = AE[v_t v_t']A' = AA'$ . Admitindo  $r_i(k) \in \mathbb{R}^m$  como sendo a resposta vetorial no horizonte  $k$  para o  $i$ -ésimo choque em uma decomposição de Cholesky de  $\Sigma$ . A função impulso-resposta  $r_a(k)$  para  $a^{(s)}$  é dada por:

$$r_a^{(s)}(k) = \sum_{i=1}^m \lambda_i r_i(k) \quad (7)$$

A equação (7), permite identificar o vetor de resposta ao impulso referente as inovações estruturais, porém, essas restrições não são condições suficientes para identificar corretamente os choques. Considerando a literatura empírica e tendo como base o que foi apresentado em Baker, Bloom e Davis (2016), os efeitos precaucionistas de um choque de incerteza se assemelham a um choque de demanda negativo. Por isso, ao determinar as restrições dos sinais adota-se o que foi definido em Redl (2017).

#### Restrição de Sinal

Ao adotar o modelo com restrições de sinais, o VAR permite identificação com maior grau de flexibilidade em relação à suposição em torno do tempo de respostas das variáveis aos choques, que resulta no conjunto de respostas identificadas para esse choque.



**Tabela 2:** Restrição de sinal para 3 até trimestres

	Selic	Horas	Consumo	Spread	IPCA	FBCF	PIB	IPE
Choque de Incerteza	-	-	?	-	?	?	?	+

Fonte: Elaboração própria.

A literatura que trata especificamente dos choques de incerteza, avalia que os efeitos desse choque na economia se equiparam aos efeitos de um choque contracionista. Nesse caso, é esperado que um aumento da incerteza econômica tenha efeitos negativos sobre a atividade econômica e sobre a inflação, como pode ser visto em Redl (2017). As restrições impostas no exercício empírico seguiram a estratégia de identificação S2 proposta por Redl (2017)<sup>35</sup>, que utilizou de um conjunto de restrições mais fracas, com duração entre 1 e 3 trimestres. O autor levanta a discussão sobre a possibilidade de a restrição de sinais oferecer identificação com mais flexibilidade sobre o tempo de resposta das variáveis ao choque, resultando numa resposta definida.

A taxa de juros Selic e a variável definida como *Spread* de crédito. Seguiu-se o que foi definido em Redl (2017) com o objetivo de controlar os efeitos temporários sobre algumas variáveis. A restrição negativa temporária imposta sobre a taxa de juros se deu unicamente como efeito indireto. Um choque de incerteza reduz a taxa de juros básica, o que conseqüentemente traz um impacto negativo no crédito e, devido à relação horas/produzividade, indiretamente, com a redução do crédito a produtividade indústria cai assim como as horas trabalhadas. Devido ao seu efeito cumulativo e considerando as expectativas formadas, a incerteza aumenta. As variáveis restantes são deixadas irrestritas tendo em vista que estamos interessados em observar o seu comportamento, principalmente após o choque de incerteza. As restrições descritas podem ser observadas na Tabela 2.

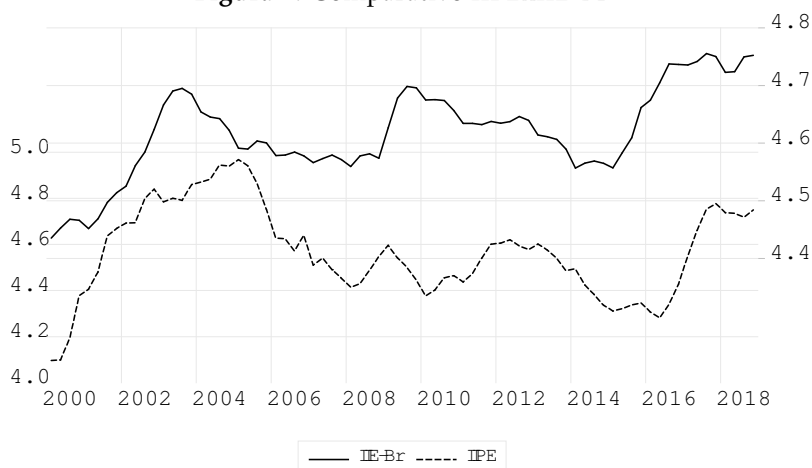
## 5 Resultados da estimação textual

Os primeiros resultados demonstram que o Índice de Incerteza da Política Econômica desenvolvido neste trabalho é capaz de capturar o sentimento de incerteza a partir das atas do Copom. Os picos apresentados pelo IPE coincidem com a realidade vivida pelo Brasil durante as duas últimas décadas, principalmente no âmbito político.

Apesar de que em teoria o IIE-Br desenvolvido por Ferreira *et al.* (2017) e do IPE construído neste trabalho serem indicadores de medidas diferentes que mensuram o grau de incerteza econômica e a incerteza da política econômica, respectivamente, optou-se por compará-los devido ao fato de serem os únicos indicadores construídos nacionalmente. Ao contrapor esses dois indicadores, observamos trajetórias divergentes e convergentes. Contudo, os períodos de ciclo e de vale apresentam o mesmo comportamento. A diferença mais significativa está no intervalo de duração e em alguns pontos na sensibilidade, conforme pode ser observado na Figura 4.

<sup>35</sup>A exceção ocorreu sobre a variável FBCF (investimento), em nosso trabalho preferimos não impor nenhuma restrição e observar o seu comportamento, enquanto que Redl (2017) impôs uma restrição negativa a variável.

Figura 4: Comparativo IPExIIE-Br



Fonte: Elaboração própria.

Nota: Índices suavizados exponencialmente.<sup>36</sup>

Enquanto o IIE-Br reduz drasticamente a incerteza após 2003 com a mudança Presidencial e a possível “euforia” dos jornais nesse ano, o IPE, por sua vez, considera a incerteza alta até 2005, resultado este influenciado pela elevação dos juros dos Estados Unidos em 2004 e pela Primeira Crise Política vivida pelo Partido dos Trabalhadores em 2005, o que justifica uma postura um pouco mais conservadora do Banco Central até 2005. Destaca-se também o fato de o IIE-Br ser mais sensível a Crise Internacional de 2008, assim como o IPE absorveu melhor a Crise na Europa. Por fim, observa-se ambos os índices adotando o mesmo curso ascendente nos três anos. Sendo assim, os primeiros resultados obtidos com o indicador construído nesse trabalho são satisfatórios, conseguiu apresentar um indicador derivado de uma metodologia mais simples e que pode auxiliar na tomada de decisão.

Em um segundo exercício comparou o IPE com o *Economic Policy Uncertainty Index Brazil*<sup>37</sup> (EPU) criado por Baker, Bloom e Davis (2016), ambos os índices estão concentrados em captar justamente esse sentimento de incerteza da política econômica mediante análise textual. A diferença se deu através da fonte de coleta e do dicionário utilizado. Enquanto o EPU Brasil utiliza jornal e um dicionário com 17 palavras o IPE tem como foco as atas das reuniões do Copom e um dicionário com 88 palavras.

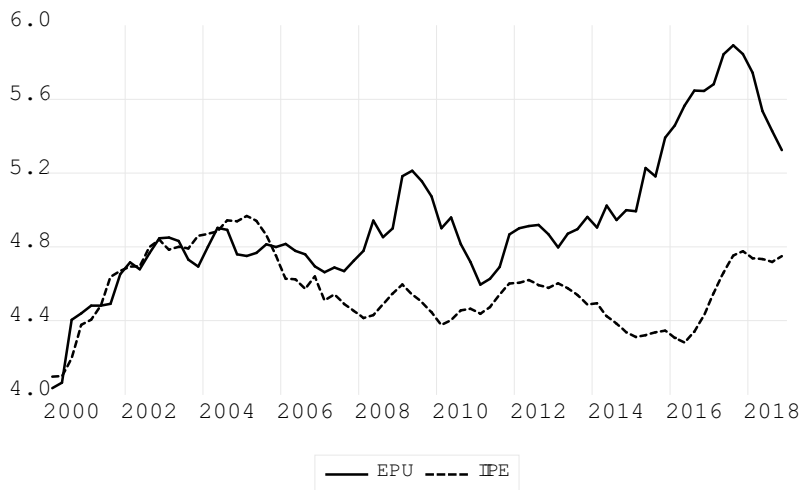
Apesar de existir grande diferença entre a quantidade de documentos utilizados por ambos indicadores<sup>38</sup>, as conclusões iniciais as trajetórias se mostraram satisfatórias. Apesar de mostrarem alguns pontos de divergências em alguns momentos, os períodos relevantes demonstraram convergência, como

<sup>36</sup>Foi utilizado o método de suavização exponencial de Holt (SEH), recomendado para realizar previsões de séries que apresentam tendência: Equação de previsão:  $\hat{y}_{t+h|t} = \ell_t + hb_t$ ; Equação de nível:  $\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$  e Equação de tendência:  $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$ . Em que:  $\ell_t$  a estimativa do nível da série no momento  $t$ ,  $b_t$  é a estimativa da tendência da série no tempo  $t$ ,  $\alpha$  é o parâmetro de suavização para o nível e  $\beta$  para a tendência.  $0 \leq \alpha, \beta \leq 1$ . A mesma fórmula foi aplicada nas séries temporais exibidas nas Figuras 5 e 7. Para maiores detalhes sobre o processo de suavização, consultar Holt (1957).

<sup>37</sup>Disponível em: <http://www.policyuncertainty.com/brazilmonthly.html>.

<sup>38</sup>O jornal possui matérias diárias, por sua vez, a ata que até 2005 era mensal, atualmente é divulgada a cada 45 dias.

Figura 5: Comparativo IPExEPU



Fonte: Elaboração própria.

Nota: Índices suavizados exponencialmente.

pôde ser observado no período de mudança de partido presidencial (2002-2003), durante a Crise na Europa (2011-2013<sup>39</sup>), e, finalmente, no período que antecedeu e sucedeu o *impeachment* da Presidente Dilma Rousseff (2016-2018), como pode ser visto na Figura 5.

Cabe ressaltar a expectativa da literatura em relação aos choques de incerteza econômica. A elevação da incerteza resulta na redução da inflação, horas trabalhadas na indústria, investimento, consumo e do PIB. Apoiado na literatura de [Costa Filho \(2014\)](#), [Baker, Bloom e Davis \(2016\)](#) e [Barboza e Zilberman \(2018\)](#), identificou-se que a influência da incerteza sobre a atividade econômica se dá por meio de canais de transmissões: *i*) consumo, *ii*) investimento e *iii*) risco.

Os canais taxas de juros e *spread* são incógnitas. Existe uma literatura que argumenta que a taxa de juros reage aos efeitos econômicos contracionistas, ocorrendo uma redução nesse indicador. Essa redução é acompanhada pela queda no *spread*. Conforme pode ser observado em [Bonciani e Van Roye \(2016\)](#) e [Meinen e Roehe \(2018\)](#), o *spread* bancário diminui em resposta a um choque de incerteza. Por sua vez, [Grimme \(2019\)](#) verifica que o *spread* ele tende a reduzir quando a incerteza exógena começa a diminuir. [Popp e Zhang \(2016\)](#) constataam que o *spread* aumentou após um choque de incerteza.

Neste trabalho, seguimos a restrição negativa imposta ao *spread* definida por [Rubio-Ramirez, Waggoner e Zha \(2010\)](#) e [Redl \(2017\)](#)<sup>40</sup>. No entanto, como exercício de robustez, retiramos a restrição negativa sobre o *spread* e os resultados observados podem ser consultados no material suplementar. Por fim, a restrição de positividade sobre o *spread* não trouxe resultados significativos.

Ao observar os efeitos da incerteza sobre o canal consumo nota-se que

<sup>39</sup>A Crise na Europa iniciou em maio de 2010, porém, os dois índices começaram a captar o aumento a incerteza doméstica em 2011.

<sup>40</sup>Estratégia de identificação S2

ela irá afetá-lo de forma negativa: as famílias reduzirão consumo ao mesmo tempo que deslocam recursos para a poupança precaucionaria; o investimento, por sua vez, é afetado de forma mais significativa, as firmas e os agentes econômicos postergam decisões à medida que o cenário de indefinição se eleva, o que ocasiona na redução da produção; o risco está ligado a possibilidade de um *default* financeiro, ou seja, esse canal sinaliza mensagens para o mercado sobre um possível colapso financeiro. Espera-se a elevação das taxas de juros como medida de “fidelização” dos agentes econômicos. Por fim, o impacto acumulado sobre a atividade econômica é negativo.

### 5.1 Choque de incerteza e seu impacto macroeconômico

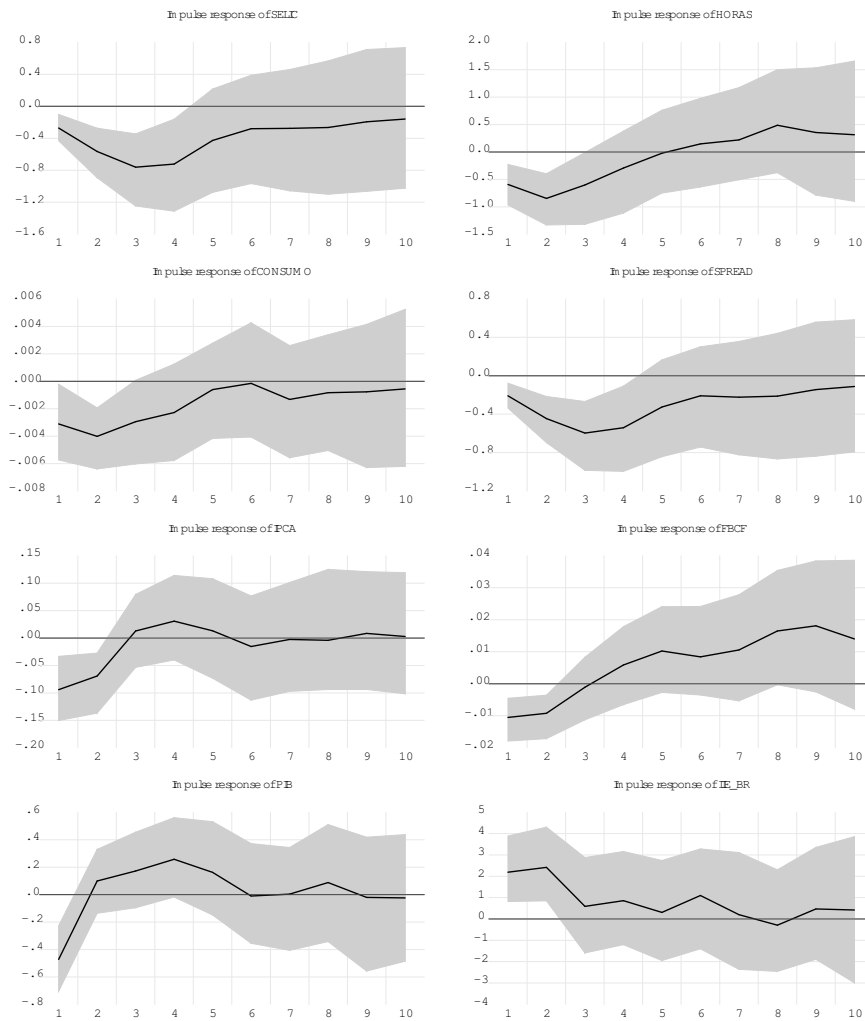
Nesta subseção, apresenta-se os resultados do modelo empírico. Aplicou-se a restrição de sinal sobre os três primeiros trimestres nas variáveis Selic, Horas, *Spread* e IPE, com objetivo de observar o impacto que um choque de incerteza causaria sobre as demais variáveis do modelo. A Figura 6 exibe as funções impulso-resposta para a restrição de sinal temporária  $k = 3$ . Dessa forma, as variáveis seguem a restrição imposta por intervalo entre o período de 1 até 3 trimestres após o choque em que: a incerteza não pode apresentar uma resposta negativa, a taxa de juros e o crédito apresentam retorno negativo e, por fim, a variável hora reduz-se durante o choque de incerteza.

A taxa Selic continua apresentando um efeito negativo significativo por um trimestre após a retirada da restrição imposta. Com relação ao *spread* do crédito, o impacto negativo do choque manteve-se significativo até o quarto trimestre. A variável horas teve o efeito imposto perdurando até a metade do segundo trimestre.

Como pode ser visto na Figura 6, as funções de impulso-resposta das variáveis irrestritas apresentaram resultados esperados, levando em consideração a literatura. Há impactos negativos e estatisticamente significativos. A inflação demonstrou uma queda inicial de 0.15 no primeiro trimestre e após o segundo, a trajetória foi de encontro até que o efeito seja nulo. Com relação ao consumo, a redução ocorre de forma mais negativa nos dois primeiros trimestres. Em seguida estabiliza-se durante um período e logo em seguida segue uma trajetória de redução dos valores negativos. O investimento evidenciou uma queda no primeiro trimestre, porém, logo em seguida apresentou um curso de redução desse efeito negativo. O PIB, num primeiro momento, apresenta nos dois primeiros trimestres uma redução, que é seguida por um leve trajeto de recuperação e novamente exibe um declínio por mais um trimestre, até que essa trajetória siga o caminho de redução dos efeitos negativos que se estende até o sexto trimestre.

Os resultados sugerem, portanto, que os setores analisados são afetados de forma significativa pela incerteza da política econômica. Trazendo esse resultado à luz do desempenho apresentado pelo Brasil ao longo dos últimos quatro anos, percebe-se a possibilidade de estabelecer uma conexão entre eles e o desempenho das variáveis macroeconômicas. Se considerarmos o fato de a incerteza apresentar um comportamento duradouro ao invés de um choque, pode-se deduzir que os efeitos sobre a atividade econômica no Brasil são consideráveis.

Figura 6: Choque de Incerteza



Fonte: Elaboração própria.

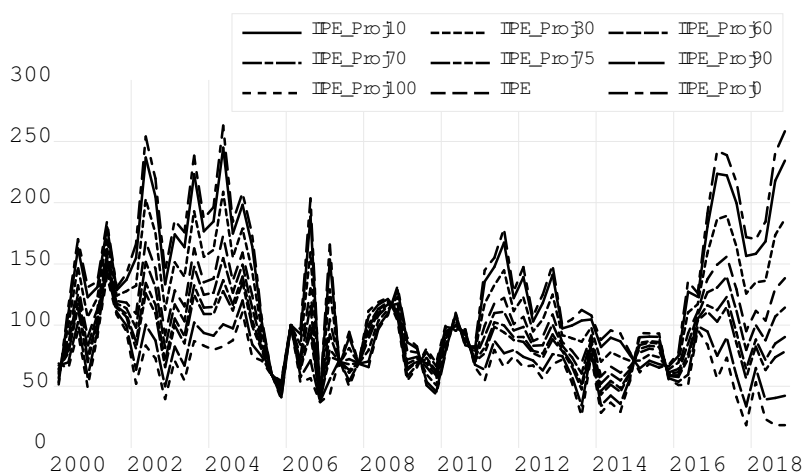
## 5.2 Robustez

O objetivo é aplicar alguns testes sobre o IPE de modo a comprovar a sua robustez.

### Indicador de Incerteza da Política Econômica

A busca por literaturas que auxiliassem a tomada de decisão em relação a ponderação do Indicador de Incerteza da Política Econômica (IPE) foi sem sucesso. Deste modo, optou-se por defini-la *ad hoc*. Decidiu-se por incorporar ao indicador final 50% do componente  $IPE_{Conjuntura}$  e 50% do  $IPE_{Projecao}$ , com objetivo de garantir a isenção da escolha, bem como também minimizar os possíveis *outliers* que cada fator poderia trazer. A Figura 7 exibe as diferentes ponderações adotadas pelo  $IPE_{Conjuntura}$  e  $IPE_{Projecao}$  na construção final do indicador. Basta observar que o resultado praticamente não muda, à

Figura 7: IIPE's



Fonte: Elaboração própria.

Nota: Todos os indicadores foram suavizados exponencialmente.

exceção quando se concentra uma parcela muito grande a um fator como é o caso do IIPE 100% projeção e do IIPE 90% conjuntura.

#### Choque de Incerteza

Tendo como objetivo testar a robustez do índice, substituímos o IIPE pelo IIE-Br e o EPU Brasil e repetimos o exercício anterior, o que permitiu analisar o impacto de um choque de incerteza advindo de outros indicadores sobre o mesmo conjunto de dados. Conforme pode ser observado nas Figuras 8 e 9, percebe-se que os resultados apresentados vão ao encontro do que foram apresentados e discutidos em 5.1. Sendo assim, considera-se robusto o Índice de Incerteza da Política Econômica construído neste trabalho.

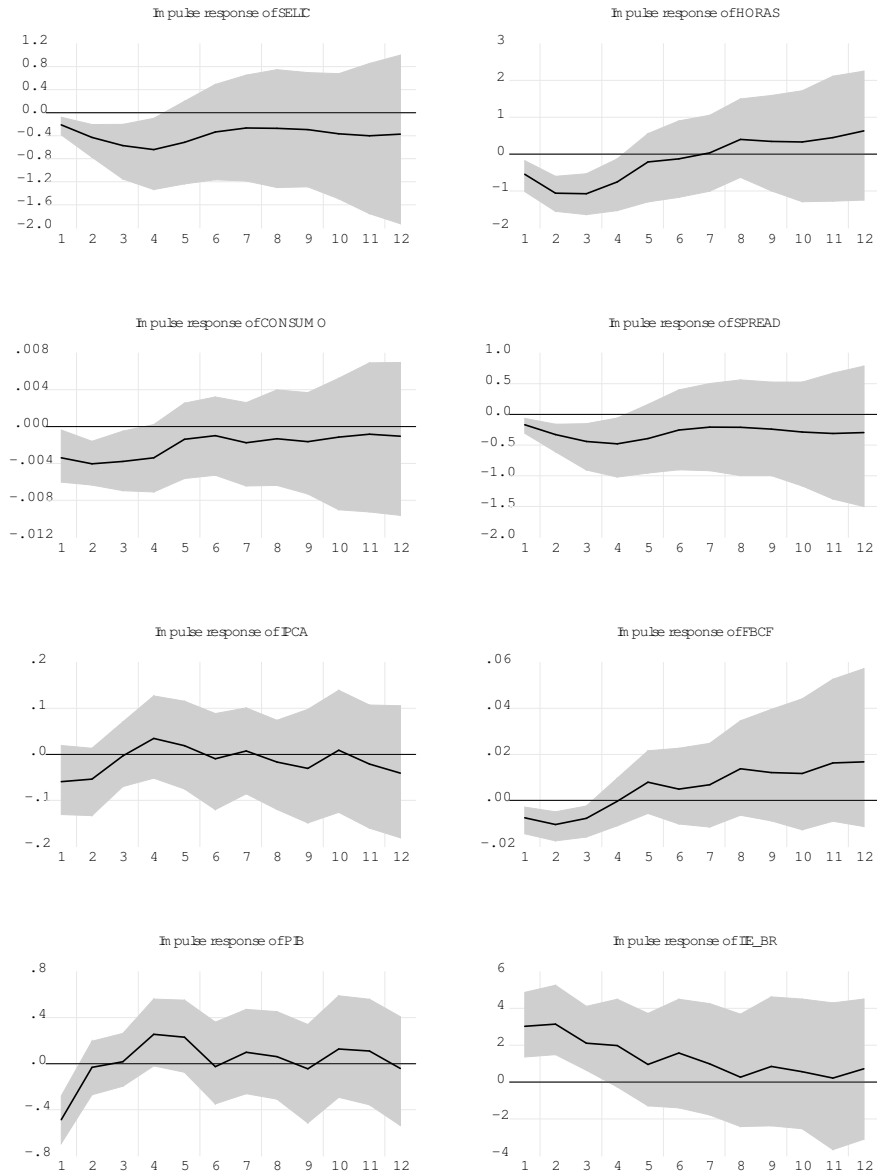
Dentro da estratégia de robustez, observou-se que os efeitos do choque de incerteza mudariam conforme a composição do IIPE. Verificou-se apenas alterações ínfimas nas funções impulso resposta, nada que modifique a análise. Os resultados desses choques podem ser vistos no Apêndice C.

Como parte final do exercício de robustez, estimou-se um modelo VAR recursivo padrão e comparou os resultados obtidos com o modelo VAR com restrição de sinais utilizado neste trabalho. A estratégia de identificação por restrição de sinais vem justamente resolver o problema da identificação recursiva. O modelo VAR padrão não consegue captar bem os efeitos dos choques de incerteza. Os resultados não correspondem com aquilo que é descrito na literatura e por isso, migrou-se para o modelo VAR com restrições de sinais. Os resultados desses choques estão disponíveis como material suplementar. As restrições de sinais, neste caso, fazem a diferença ao captar melhor o efeito de um choque de incerteza.

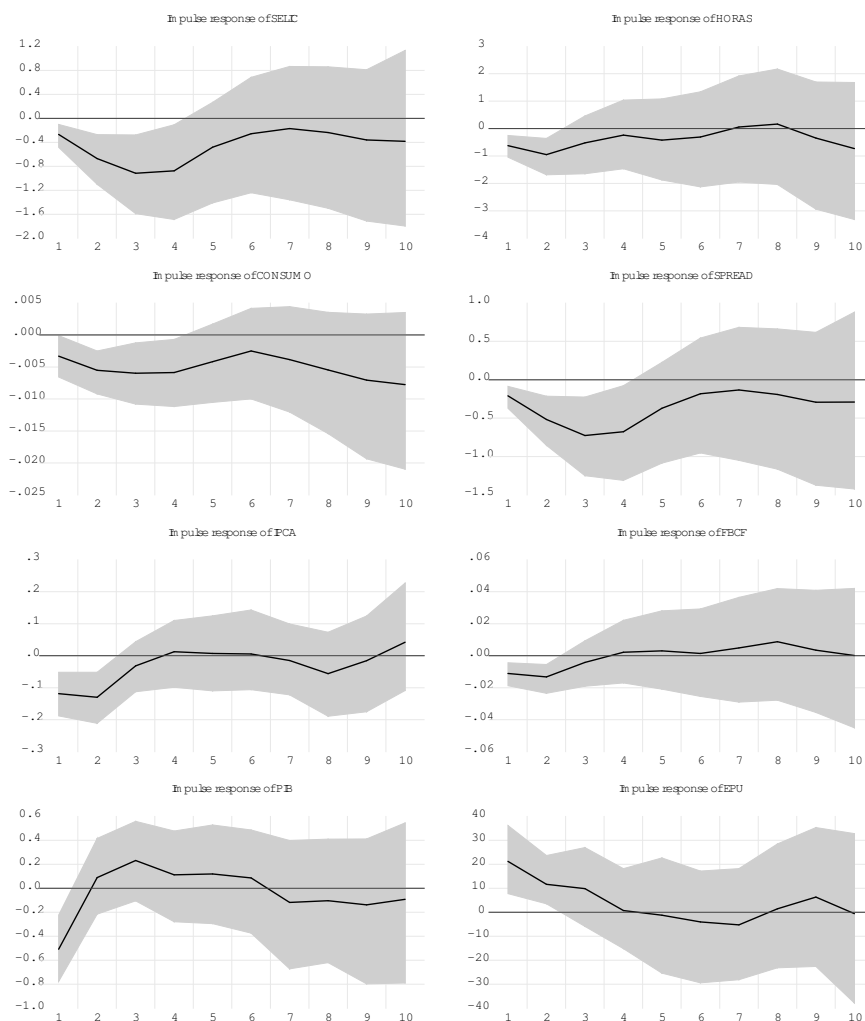
## 6 Considerações finais

Os primeiros resultados demonstraram que o Índice de Incerteza da Política Econômica desenvolvido nesse estudo é capaz de capturar o sentimento de incerteza unicamente a partir das atas do Copom. Ao realizar um comparativo

**Figura 8: Choque de Incerteza: IIE-Br**



Fonte: Elaboração própria.

**Figura 9:** Choque de Incerteza: EPU Brasil

Fonte: Elaboração própria.

com os demais indicadores, a conclusão em um primeiro momento mostrou ser satisfatória. Sabe-se que é possível buscar novas formas de refinar ainda mais o índice construído neste estudo, tendo como objetivo principal aperfeiçoar sua precisão na captura desse sentimento.

Um ponto que merece destaque é a eliminação de um possível viés do intérprete<sup>41</sup>: os indicadores estimados por meio de artigos de jornais, podem muitas vezes refletir uma análise própria do jornalista ao interpretar as informações divulgadas nas atas das reuniões, além do uso de alguns termos que podem passar uma sensação de incerteza ao leitor<sup>42</sup>.

Espera-se a existência de um viés de escrita por parte de quem divulga

<sup>41</sup>Ferreira *et al.* (2017) justifica que ao utilizar apenas um jornal na sua análise o resultado poderia conter viés.

<sup>42</sup>Alguns autores podem ter preferência por algumas palavras que via dicionário podem exprimir um sentimento diferente.



a ata, até como medida de reduzir o impacto sobre as expectativas do mercado. Por isso, ao optar por gerar o índice diretamente dos documentos oficiais, retira-se esse provável segundo viés.

Em contrapartida, a percepção de incerteza da política econômica por parte do Copom tende a mudar conforme a composição dos seus membros se altera. Em outras palavras, essa percepção não será homogênea ao longo do tempo, visto que tende a refletir as diferentes visões dos membros do Copom, que mudam ao longo do tempo. No entanto, este resultado já era esperado e a métrica aplicada neste trabalho “pune” os documentos que apresentam termos *outliers*, atribuindo um peso maior para aqueles que seguem uma distribuição homogênea ao longo de toda amostra analisada e, reduzindo assim, os efeitos da heterogeneidade sobre o sentimento de incerteza.

Ao se trabalhar com o indicador ponderado pela atual conjuntura e por aquilo que os formuladores de política econômica projetam, conseguiu-se melhorar a capacidade do índice de captar melhor o sentimento de incerteza, tendo em vista que ambos os fatores estão de certa forma interligados.

Num primeiro momento, observa-se as variáveis em que foram aplicadas a restrição de sinais durante o período de três trimestres, a taxa Selic e o *Spread* de crédito continuam apresentando uma resposta negativa ao aumento da incerteza. Com relação às variáveis em que não foram aplicadas nenhuma restrição de sinal, um choque de 1% na incerteza produz efeitos significativos sobre os setores analisados. Imediatamente, o consumo é reduzido e os consumidores adotam uma postura precaucionária em relação ao consumo. Esse desaquecimento do consumo combinado com a elevação da incerteza reflete na redução da inflação, como pode ser visto na Figura 6. Como esperado, o Investimento também é afetado negativamente. A literatura econômica aponta que sob esse cenário é inevitável a reorganização do portfólio de investimento. Existe uma tendência de buscar ativos com segurança garantida. O PIB, que dada a junção dos resultados descritos acima unido ao choque demonstrou um efeito negativo que perdura até 6 trimestres.

A incerteza possui um comportamento contracíclico e retroalimentador. Choques de incerteza terão um efeito negativo sobre a atividade econômica. Os resultados alcançados neste trabalho estão de acordo com o que foi apresentado por Costa Filho (2014) e Barboza e Zilberman (2018). Porém, fez-se o uso de um conjunto maior de variáveis macroeconômicas e de apenas uma variável de incerteza. Conclui-se que a incerteza influencia a postura adotada por empresários que postergam decisões de produção e investimento e agentes econômicos que buscam uma reorganização de portfólio com objetivo de garantir a segurança de seus ativos.

É fundamental chamar atenção para o fato de que a incerteza agora é uma variável mensurável e que pode ser utilizada para justificar e entender o comportamento tanto dos agentes econômicos quanto dos agentes políticos. É possível estabelecer um elo entre os resultados advindos deste estudo e o desempenho macroeconômico do Brasil nos últimos 4 anos. Por fim, O Indicador de Incerteza da Política Econômica (IIFE) criado neste trabalho pode auxiliar tanto as Autoridades Monetárias quanto os agentes econômicos a respeito da tomada de decisões, além de ser uma boa *proxy* para estudos futuros dentro deste cenário.

## 7 Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Os autores também agradecem aos comentários pertinentes do revisor anônimo.

## Referências

- AGUIAR, E J de *et al.* Análise de sentimento em redes sociais para a língua portuguesa utilizando algoritmos de classificação. *In: SBC. ANAIS do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos.* [S. l.: s. n.], 2018.
- ALCHIAN, A A. Uncertainty, evolution, and economic theory. *Journal of Political Economy*, v. 58, n. 3, p. 211–221, 1950.
- ALUÍSIO, S M; ALMEIDA, G M B. O que é e como se constrói um corpus? Lições aprendidas na compilação de vários corpora para pesquisa linguística. *Calidoscópio*, v. 4, n. 3, p. 156–178, 2006.
- APEL, M; GRIMALDI, M. The information content of central bank minutes, 2012.
- BAKER, S R; BLOOM, N; DAVIS, S J. Measuring economic policy uncertainty. *Quarterly Journal of Economics*, v. 131, n. 4, p. 1593–1636, 2016.
- BARBOZA, R de M; ZILBERMAN, E. Os efeitos da incerteza sobre a atividade econômica no Brasil. *Revista Brasileira de Economia*, v. 72, n. 2, p. 144–160, 2018.
- BHOLAT, D *et al.* Text mining for central banks. *Available at SSRN 2624811*, 2015.
- BONCIANI, D; VAN ROYE, B. Uncertainty shocks, banking frictions and economic activity. *Journal of Economic Dynamics and Control*, v. 73, p. 200–219, 2016.
- BRUNO, G. Central Bank Communications: Information extraction and Semantic Analysis. *In: THE R User Conference, useR! 2017 July 4-7 2017 Brussels, Belgium.* [S. l.: s. n.], 2017. p. 253.
- BRUNO, G. Text mining and sentiment extraction in central bank documents. *In: IEEE. 2016 IEEE International Conference on Big Data (Big Data).* [S. l.: s. n.], 2016. p. 1700–1708.
- CHISHOLM, E; KOLDA, T G. *New term weighting formulas for the vector space method in information retrieval.* [S. l.], 1999.
- CORREA, R *et al.* Sentiment in Central Banks' Financial Stability Reports. *Available at SSRN 3091943*, 2017.

- COSTA FILHO, A E. Incerteza e atividade econômica no Brasil. *Economia Aplicada*, v. 18, n. 3, p. 421–453, 2014.
- COSTA FILHO, A E; ROCHA, F. Como o mercado de juros futuros reage à comunicação do Banco Central? *Economia Aplicada*, v. 14, n. 3, p. 265–292, 2010.
- COSTA FILHO, A E; ROCHA, F. Comunicação e política monetária no Brasil. *Revista Brasileira de Economia*, v. 63, n. 4, p. 405–422, 2009.
- FERREIRA, P C *et al.* Medindo a Incerteza Econômica no Brasil. *Economia Aplicada*, 2017.
- GODEIRO, L L; OLIVEIRA LIMA, L R R de. Medindo incerteza macroeconômica para o Brasil. *Economia Aplicada*, v. 21, n. 2, p. 311, 2017.
- GRAMINHO, F M *et al.* Sentimento e Macroeconomia: uma análise dos índices de confiança no Brasil. *Brazilian Central Bank, trabalhos para discussão*, v. 408, 2015.
- GRIGNANI, F; FONTANA, R. *Using machine learning and Bayesian networks to objectively analyze central bank statements and market sentiment*. 2018. Tese (Doutorado) – Politecnico di Torino.
- GRIMME, C. Uncertainty and the cost of bank vs. bond finance. CESifo Working Paper, 2019.
- HANSEN, S; MCMAHON, M. Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, v. 99, s114–s133, 2016.
- HASELMAYER, M; JENNY, M. Sentiment analysis of political communication: combining a dictionary approach with crowdcoding. *Quality & quantity*, v. 51, n. 6, p. 2623–2646, 2017.
- HOLT, C C. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, v. 20, n. 1, p. 5–10, 1957.
- HUTTO, C J; GILBERT, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: EIGHTH international AAAI conference on weblogs and social media. [S. l.: s. n.], 2014.
- JURADO, K; LUDVIGSON, S C; NG, S. Measuring uncertainty. *American Economic Review*, v. 105, n. 3, p. 1177–1216, 2015.
- KEARNEY, C; LIU, S. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, v. 33, p. 171–185, 2014.
- KNIGHT, F H. *Risk, uncertainty and profit*. [S. l.]: Houghton Mifflin Company, New York, 1921.

- LI, J *et al.* Topic Popularity Prediction with Sentiment Time Series on Short Text based Social Media, 2019.
- LOUGHRAN, T; MCDONALD, B. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, v. 66, n. 1, p. 35–65, 2011.
- MACHADO, M A V *et al.* Análise do Sentimento Textual dos Relatórios de Desempenho Trimestral das Indústrias Brasileiras. *Sociedade, Contabilidade e Gestão*, v. 12, n. 1, 2017.
- MEINEN, P; ROEHE, O. To sign or not to sign? On the response of prices to financial and uncertainty shocks. *Economics Letters*, v. 171, p. 189–192, 2018.
- MONTES, G C; NICOLAY, R T F. Central bank's perception on inflation and inflation expectations of experts: empirical evidence from Brazil. *Journal of Economic Studies*, v. 42, n. 6, p. 1142–1158, 2015.
- NOPP, C; HANBURY, A. Detecting Risks in the Banking System by Sentiment Analysis. In: PROCEEDINGS of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal: Association for Computational Linguistics, set. 2015. p. 591–600. Disponível em: <https://aclweb.org/anthology/D/D15/D15-1071>.
- ORMEROD, P; NYMAN, R; TUCKETT, D. Measuring financial sentiment to predict financial instability: A new approach based on text analysis. *arXiv preprint arXiv:1508.05357*, 2015.
- PAO, M L. Automatic text analysis based on transition phenomena of word occurrences. *Journal of the American Society for Information Science*, v. 29, n. 3, p. 121–124, 1978.
- POPP, A; ZHANG, F. The macroeconomic effects of uncertainty shocks: The role of the financial channel. *Journal of Economic Dynamics and Control*, v. 69, p. 319–349, 2016.
- REDL, C. The impact of uncertainty shocks in the United Kingdom. Bank of England Working Paper, 2017.
- RUBIO-RAMIREZ, J F; WAGGONER, D F; ZHA, T. Structural vector autoregressions: Theory of identification and algorithms for inference. *The Review of Economic Studies*, v. 77, n. 2, p. 665–696, 2010.
- RYBINSKI, K I. A machine learning framework for automated analysis of formal and informal central bank communication. The case of the National Bank of Poland, 2018.
- SALTON, G; BUCKLEY, C. Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, v. 24, n. 5, p. 513–523, 1988.
- SCHYMURA, L G. A crescente importância de medir a incerteza e seus impactos no Brasil de hoje. *Revista Conjuntura Econômica*, v. 71, n. 5, p. 6–9, 2017.

- SILVA, M D O P *et al.* O efeito do sentimento das notícias sobre o comportamento dos preços no mercado acionário brasileiro, 2017.
- UHLIG, H. What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics*, v. 52, n. 2, p. 381–419, 2005.
- WILKINSON, L; FRIENDLY, M. The history of the cluster heat map. *The American Statistician*, v. 63, n. 2, p. 179–184, 2009.
- YOU, S; DESARMO, J; JOO, S. Measuring happiness of US cities by mining user-generated text in Flickr. com: A pilot analysis. *In: AMERICAN SOCIETY FOR INFORMATION SCIENCE. PROCEEDINGS of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries. [S. l.: s. n.], 2013. p. 167.*

## **Apêndice A Lista de palavras**

### **A.1 Incerteza da Política Econômica**

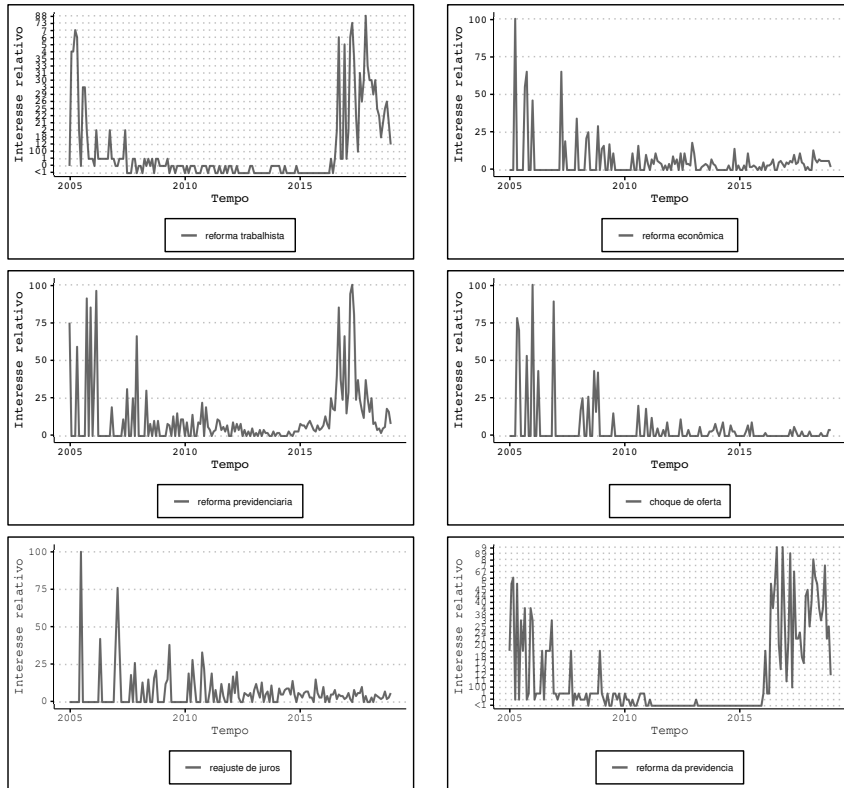
congress; deficit; deficits; economic; economy; instabilities; instability; readjustment; readjustments; reform; reforms; regulation; regulations; risk; risked; riskier; riskiest; riskiness; risking; risks; risky; rumors; senate; shock; shocks; speculate; speculated; speculates; speculating; speculation; speculations; speculative; speculatively; uncertain; uncertainly; uncertainties; uncertainty; unexpected; unexpectedly; unpredictable; unpredictably; unpredicted; volatile; volatilities; volatility.

### **A.2 Projeção de Incerteza da Política Econômica**

alteration; alterations; anomalies; anomaly; anticipate; assume; assumption; believes; cautiousness; doubts; imprecise; improbable; indefinite; indeterminate; maybe; might; ordinarily; perhaps; possibilities; possibility; possible; precaution; precautionary; precautions; predict; predicted; prediction; predictive; predictor; presumes; presuming; presumption; probable; probably; rumors; speculate; speculated; speculates; speculation; speculations; suggest; suggested; suggests.

## Apêndice B Pesquisas Google Trends Brasil

Figura B.1: Termos Próprios Google Trends Brasil

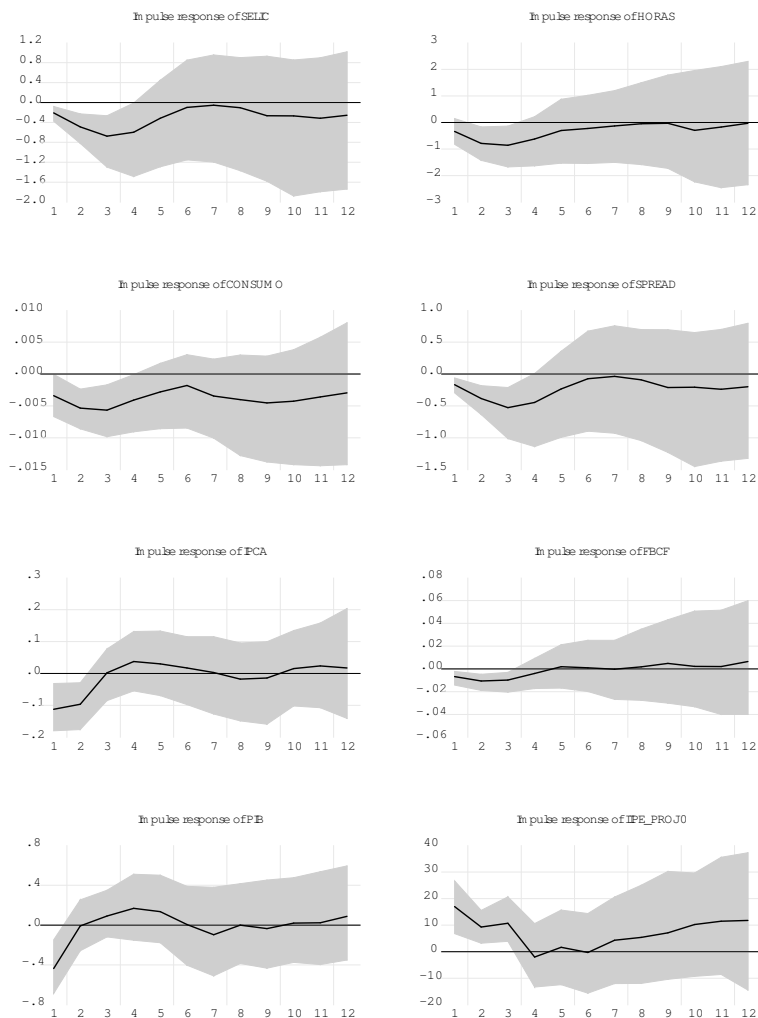


Fonte: Elaboração Própria.

Nota: Os Termos são tratados como bigramas.

### Apêndice C Choques de Incerteza: diferentes composições do IIPE

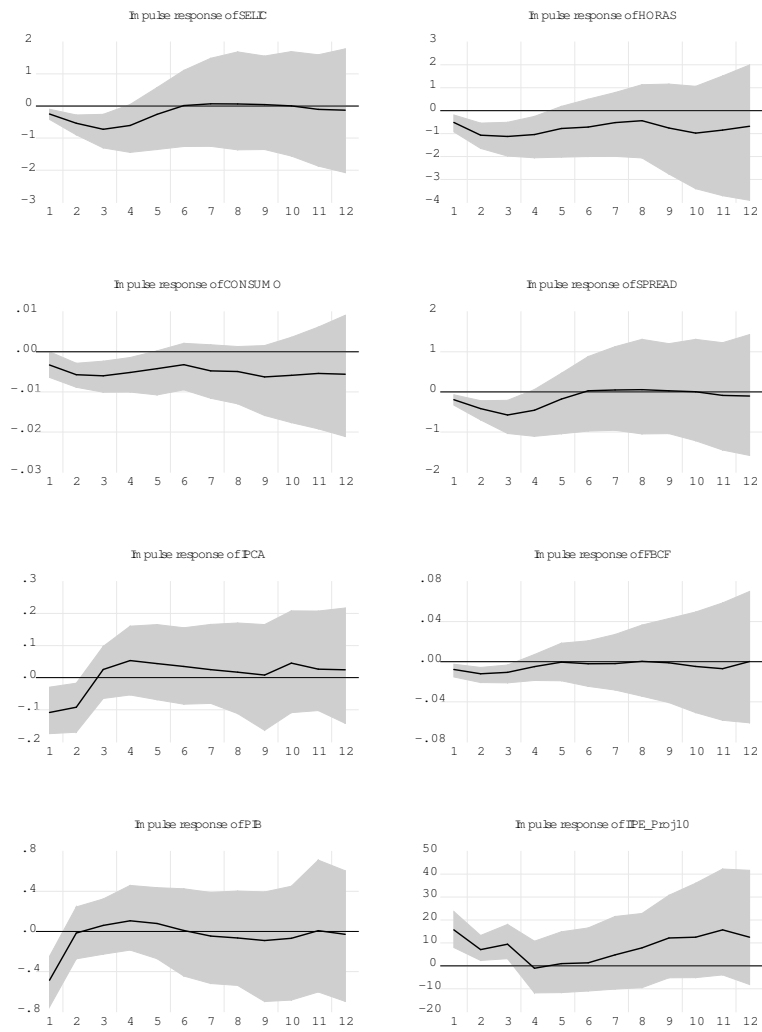
Figura C.1: Choque de Incerteza: IIPE\_Proj 0



Fonte: Elaboração Própria

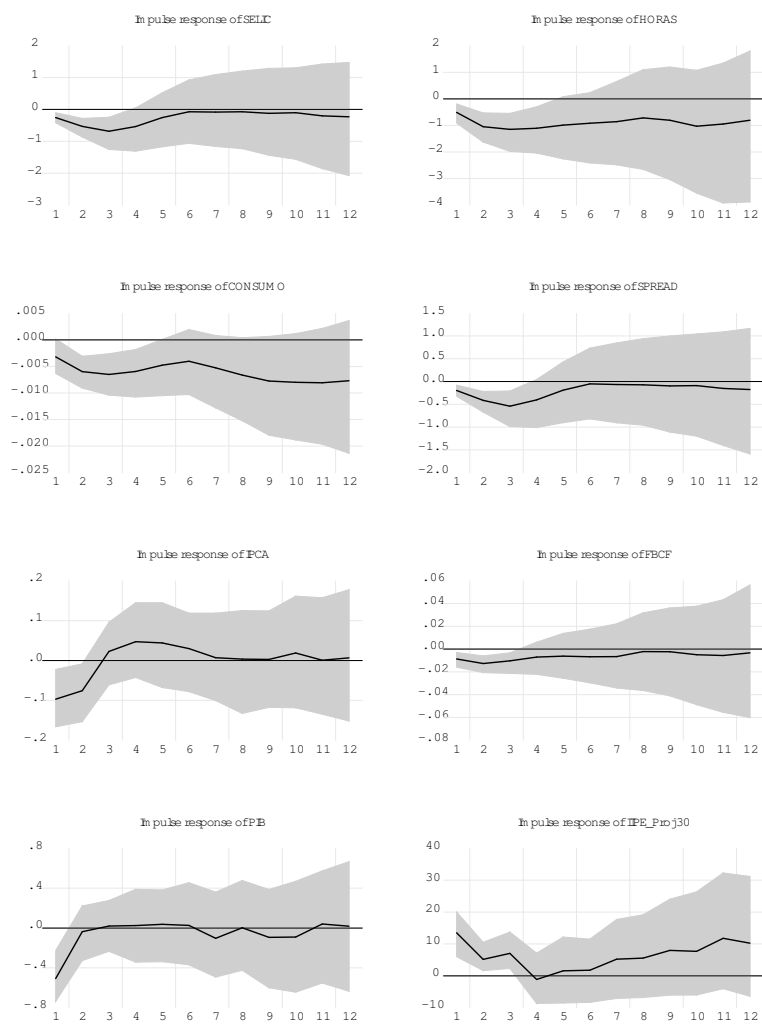


**Figura C.2:** Choque de Incerteza: IIFE\_Proj 10



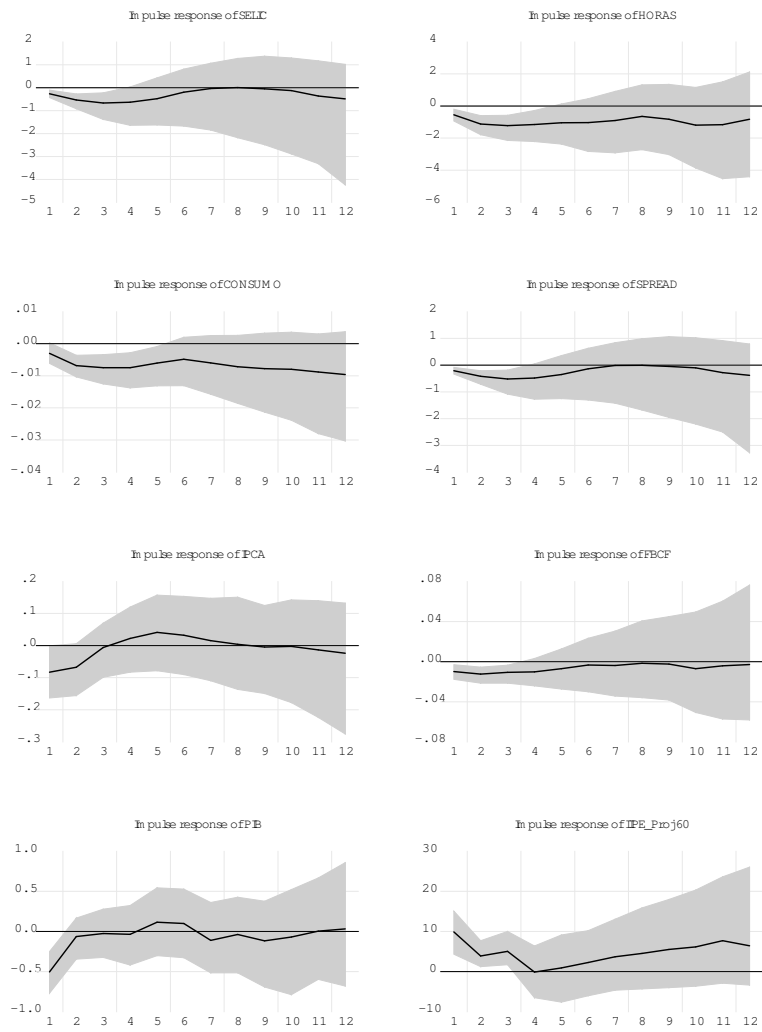
Fonte: Elaboração Própria

Figura C.3: Choque de Incerteza: IIPE\_Proj 30



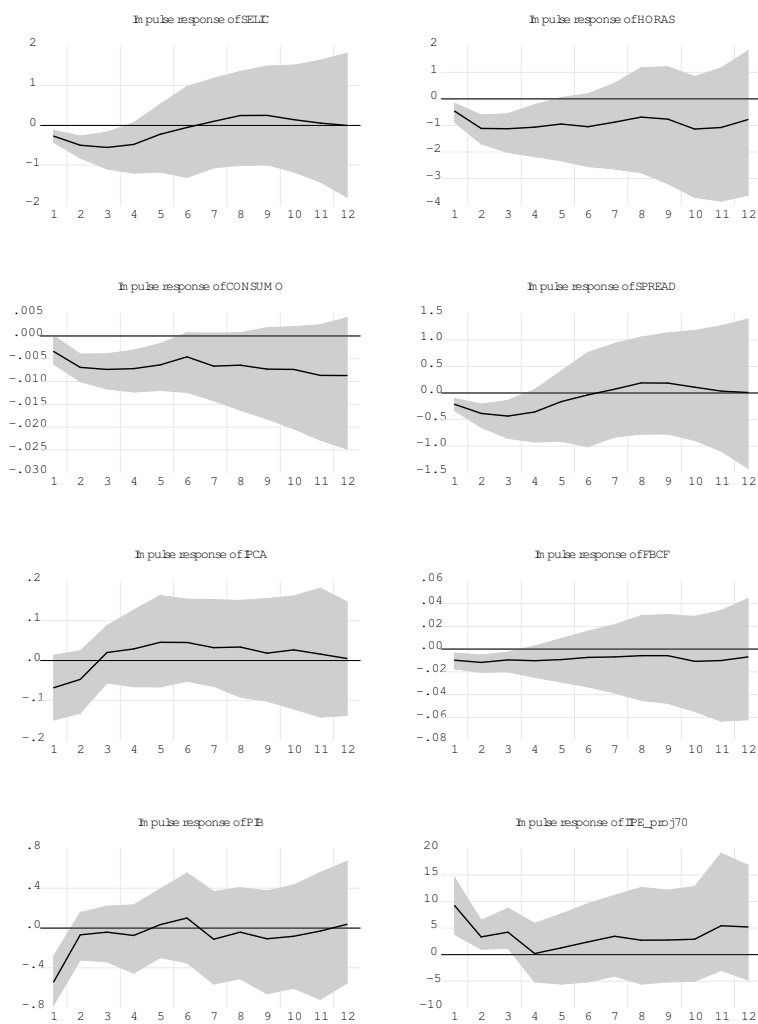
Fonte: Elaboração Própria

Figura C.4: Choque de Incerteza: IIFE\_Proj 60



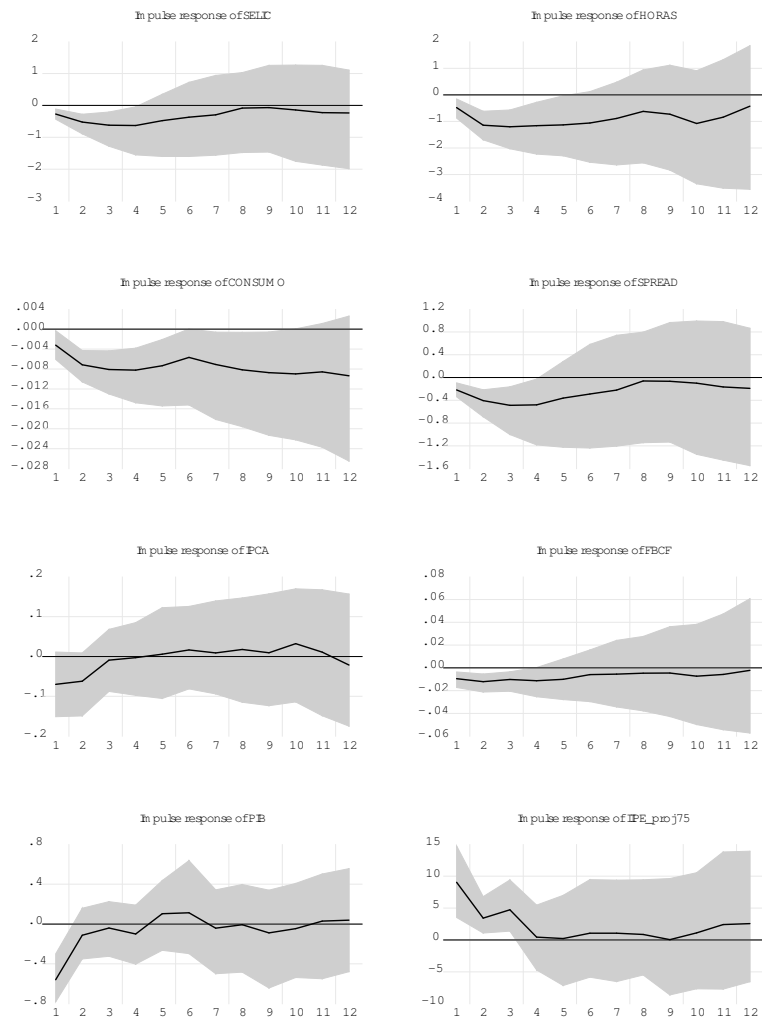
Fonte: Elaboração Própria

**Figura C.5: Choque de Incerteza: IIPE\_Proj 70**



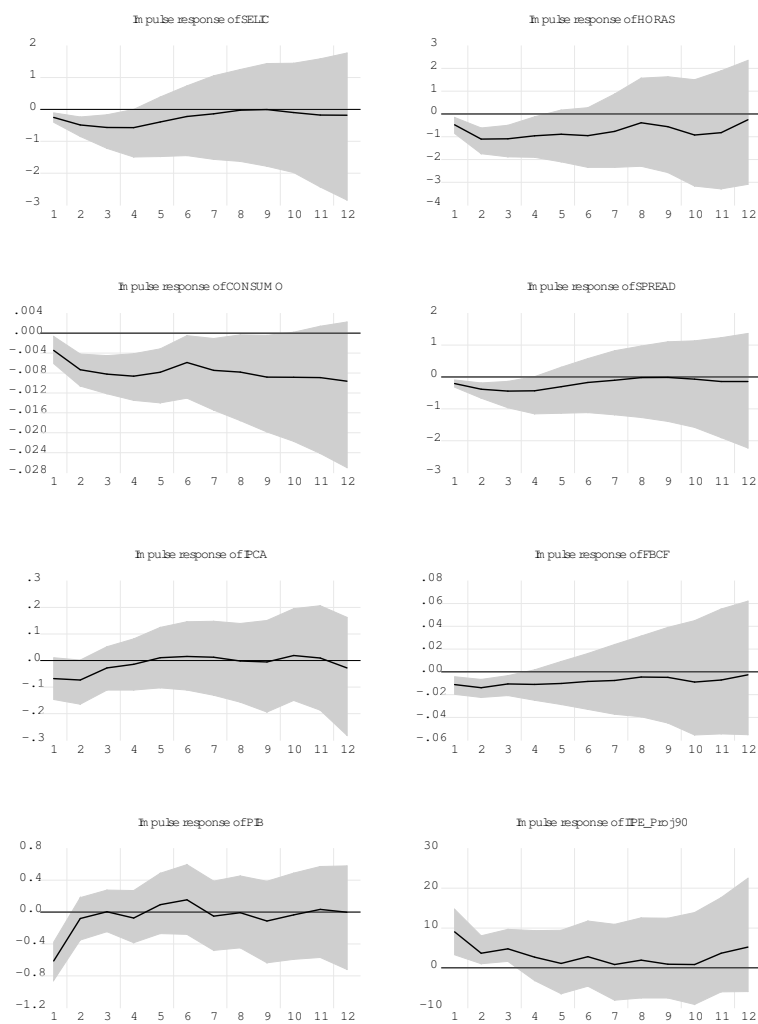
Fonte: Elaboração Própria

Figura C.6: Choque de Incerteza: IIPE\_Proj 75



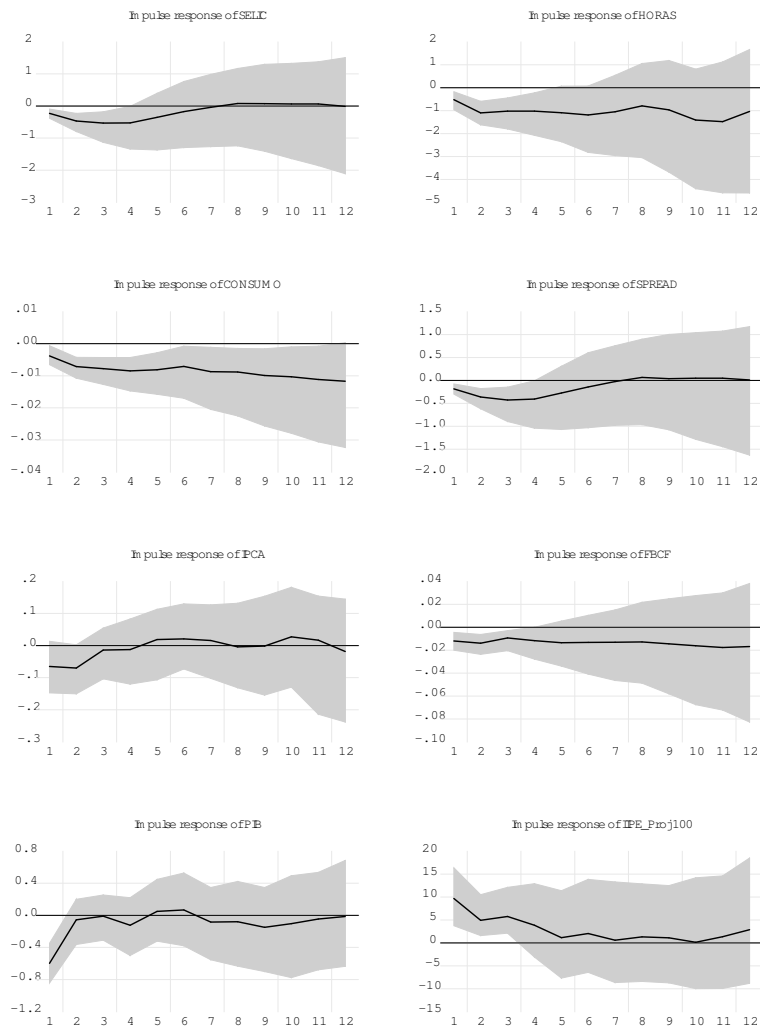
Fonte: Elaboração Própria

Figura C.7: Choque de Incerteza: IIPE\_Proj 90



Fonte: Elaboração Própria

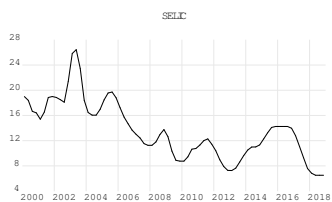
**Figura C.8:** Choque de Incerteza: IIFE\_Proj 100



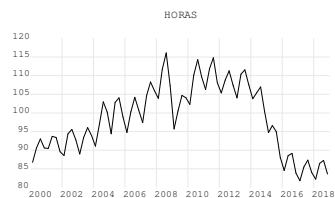
Fonte: Elaboração Própria

## Apêndice D Séries temporais

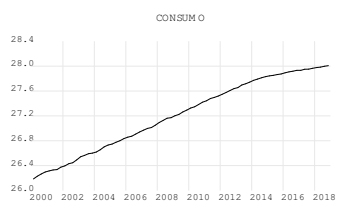
Figura D.1: Séries temporais



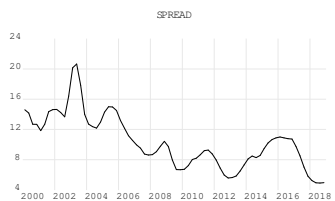
(a) Selic



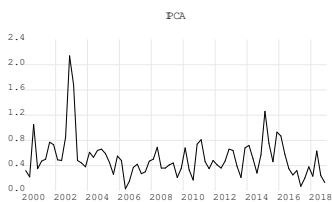
(b) Horas trabalhadas na indústria



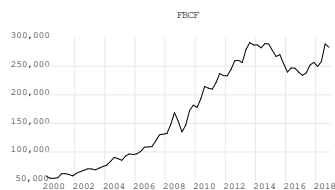
(c) Consumo total (em log)



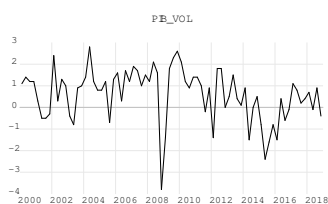
(d) Spread



(e) IPCA



(f) FBCF



(g) PIB