



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

Câmpus de Presidente Prudente

Roger Arthur Novaes Peron

Modelagem econométrica dos principais indicadores
econômicos do Brasil: pré e pós a crise financeira mundial de
2008

PRESIDENTE PRUDENTE

2023

Roger Arthur Novaes Peron

Modelagem econométrica dos principais indicadores econômicos do Brasil: pré e pós a crise financeira mundial de 2008

Relatório Final de Trabalho de Conclusão de Curso
apresentado ao Curso de Graduação em Estatística
da FCT/Unesp para aproveitamento na disciplina
de Trabalho de Conclusão de Curso.

Orientador: Prof. Dr. Manoel Ivanildo Silvestre
Bezerra.

PRESIDENTE PRUDENTE

2023

P453m

Peron, Roger Arthur Novaes

Modelagem econométrica dos principais indicadores econômicos do Brasil: pré e pós a crise financeira mundial de 2008 / Roger Arthur Novaes Peron. -- Presidente Prudente, 2023

74 p.

Trabalho de conclusão de curso (Bacharelado - Estatística) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente

Orientador: Manoel Ivanildo Silvestre Bezerra

1. crise financeira mundial de 2008. 2. modelos decomposição de séries temporais. 3. variáveis dummy. 4. matriz de delineamento. 5. harmônicas. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

TERMO DE APROVAÇÃO

Roger Arthur Novaes Peron

Modelagem econométrica dos principais indicadores econômicos do Brasil: pré e pós a crise financeira mundial de 2008

Relatório Final de Trabalho de Conclusão de Curso aprovado como requisito para obtenção de créditos na disciplina Trabalho de Conclusão do curso de graduação em Estatística da Faculdade de Ciências e Tecnologia da Unesp, pela seguinte banca examinadora:

Orientador:



Prof. Dr. Manoel Ivanildo Silvestre Bezerra
Departamento de Estatística



Prof. Dr. Fernando Antonio Moala
Departamento de Estatística



Prof. Dr. Jose Gilbeto Spasiani Rinaldi
Departamento de Estatística

Presidente Prudente, 05 de dezembro de 2023.

RESUMO

A crise financeira mundial de 2008 foi a maior crise desde 1929, vários países do mundo sentiram seu impacto e entre eles está o Brasil. Uma forma de medir esse impacto na economia é através de alguns indicadores econômicos, e esse trabalho tem como objetivo utilizar técnicas estatísticas para modelar alguns desses indicadores como o IPCA (Índice de Preço ao Consumidor Amplo), Preços das Exportações e Importações e as Taxas de Juros e de Câmbio.

Para as análises foram coletados valores mensais dos anos de 2006 a 2009, foram aplicados os modelos de regressão de decomposição temporal e de regressão harmônico (senos e cossenos) junto com variáveis “*dummy*” Para verificar se houve diferença entre os períodos pré e pós crise foram definidas as variáveis “*dummy*” Z_{it} , sendo -1 para o período pré-crise (2006-2008) e 1 para o pós-crise (2008-2009), e $Z_{it}t$ para verificar se houve diferença na inclinação da reta entre esses períodos, ou seja, se houve mudança na tendência dos indicadores e no deslocamento da reta (intercepto).

Para a análise residual foi utilizado os testes de Shapiro-Wilk e Durbin-Watson e análises gráficas, já para verificar a adequabilidade dos modelos foram utilizados a raiz do erro quadrático médio (RMSE), o erro médio absoluto (MAE) e o coeficiente de determinação (R^2).

Os modelos obtiveram bons resultados de maneira geral, todos apontaram diferença significativa entre o período pré e pós crise, entretanto os resíduos do modelo utilizando harmônicos para o IPCA não passaram no teste de Durbin-Watson e para o Índice de Preços das Importações e Exportações o teste foi inconclusivo.

Os modelos utilizando a matriz de delineamento não tiveram problemas com autocorrelação dos resíduos e não tiveram um ajuste muito inferior se comparado aos harmônicos, entretanto apenas a Taxa Juros obteve um parâmetro significativo para a componente sazonal, logo os modelos se reduziram a uma regressão linear múltipla com variáveis ‘*dummy*’, pois apenas a componente de tendência foi significativa.

Palavras-chave: crise financeira mundial de 2008, indicadores econômicos, modelos decomposição de séries temporais, variáveis *dummy*, matriz de delineamento e harmônicas.

ABSTRACT

The 2008 global financial crisis was the most significant crisis since 1929, impacting various countries worldwide, including Brazil. One way to measure this impact on the economy is through several economic indicators. This study aims to use statistical techniques to model some of these indicators, such as the IPCA (Consumer Price Index), Export and Import Prices, and Interest and Exchange Rates. Monthly values from 2006 to 2009 were collected for the analyses. Regression models, including temporal decomposition and harmonic regression (sinusoidal and cosinusoidal functions), along with dummy variables, were applied.

To assess the difference between pre and post-crisis periods, dummy variables Z_{it} were defined, with -1 representing the pre-crisis period (2006-2008) and 1 representing the post-crisis period (2008-2009). Additionally, $Z_{it}t$ was defined to examine if there was a difference in the slope of the line between these periods, indicating a change in the indicators' trends and intercept displacement.

For residual analysis, Shapiro-Wilk and Durbin-Watson tests, along with graphical analyses, were employed. To assess model adequacy, root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2) were used.

In general, the models yielded good results, indicating a significant difference between the pre and post-crisis periods. However, the residuals of the harmonic model for IPCA did not pass the Durbin-Watson test, and the test for the Import and Export Price Index was inconclusive.

The models using the design matrix did not encounter issues with autocorrelation of residuals and did not exhibit a significantly inferior fit compared to the harmonic models. However, only the Interest Rate obtained a significant parameter for the seasonal component. Consequently, the models were simplified to a multiple linear regression with dummy variables, as only the trend component proved to be statistically significant.

Keywords: Global Financial Crisis of 2008, Economic Indicators, Time Series Decomposition Models, Dummy Variables, Design Matrix, Harmonics.

SUMÁRIO

1 INTRODUÇÃO	8
2 METODOLOGIA.....	10
2.1 Modelo de Regressão Linear Simples	11
2.2 Modelo de Regressão Linear Múltipla	14
2.3 Decomposição de Modelos de Séries Temporais.....	16
2.3.1 Teste de Dickey-Fuller	18
2.3.1 Modelos de Decomposição com Variáveis Dummy	19
2.3.2 Modelos de Decomposição com Harmônicas	20
2.3.3 Modelos de Decomposição Temporal	22
2.3.4 Mínimos Quadrados Generalizados	23
2.4 Análise de Resíduos.....	25
2.4.1 Teste de Shapiro-Wilk.....	25
2.4.2 Teste de Durbin-Watson	26
2.4.3 Análises Gráficas	27
2.5 Medidas para o Diagnóstico dos Modelos	27
3 RESULTADOS.....	28
3.1 Análises Descritivas dos Indicadores	28
3.2 Modelagem e Análise dos Indicadores	34
3.2.1 Modelagem da Taxa de Juros Over/Selic	34
3.2.2 Análise dos Resíduos e Adequabilidade dos Modelos da Taxa de Juros Over/Selic	39

3.2.3 Modelagem da Taxa de Câmbio	41
3.2.4 Análise dos Resíduos e Adequabilidade dos Modelos da Taxa de Câmbio ...	44
3.2.5 Modelagem do IPCA.....	47
3.2.6 Análise dos Resíduos e Adequabilidade dos Modelos do IPCA.....	50
3.2.7 Modelagem do Índice de Preços das Importações.....	53
3.2.6 Análise dos Resíduos e Adequabilidade dos Modelos do Índice de Preços das Importações	56
3.2.7 Modelagem do Índice de Preços das Exportações.....	59
3.2.6 Análise dos Resíduos e Adequabilidade dos Modelos do Índice de Preços das Exportações.	63
4 CONCLUSÕES	66
REFERÊNCIAS	67
APÊNDICE	68

1 INTRODUÇÃO

Crises econômicas são comuns na história recente e isso é algo esperado visto que a economia é cíclica, ou seja, em alguns momentos há crescimento já em outros há queda econômica, o que é normal dentro do sistema capitalista. Podemos dizer que crises são esperadas, entretanto o que muda são suas razões e a intensidade com que essa crise irá afetar os países, e para medir os efeitos existem vários indicadores econômicos que ajudam a mensurar e indicar os problemas, auxiliando o poder pública a tomar decisões e medidas para tentar reverter essa situação.

Existem algumas instituições e fundações que montam e divulgam esses indicadores, como por exemplo o Instituto Brasileiro de Geografia e Estatística (IBGE), que é responsável por indicadores como o Índice Nacional de Preços ao Consumidor Amplo (IPCA) e o Índice Nacional de Preços ao Consumidor (INPC). Temos a Fundação Getúlio Vargas (FGV) que organiza o Índice Geral de Preços de Mercado (IGP-M) e o Índice Geral de Preços/Disponibilidade Interna (IGP-DI), existe também a Fundação Instituto de Pesquisas Econômicas (FIPE) que realiza o Índice de Preços ao Consumidor (IPC). Existem outros índices econômicos que também são usados como a Taxa de Desemprego e o Produto Interno Bruto (PIB).

A quebra em 15 de setembro de 2008 do Lehman Brothers, segundo maior banco de investimentos dos EUA na época foi o estopim da crise financeira de 2008, também conhecida como crise do *subprime*, a maior crise global desde a grande depressão, a partir da quebra da bolsa de Nova York em 1929. Na década de 90 os preços dos imóveis estavam subindo continuamente, formando uma bolha imobiliária que nada mais é que inflar artificialmente os preços através de uma grande especulação e gerando assim uma falsa “prosperidade”, e essa bolha veio a “estourar” em 2008 acarretando a crise. Naquela época o governo garantia a liquidez do mercado através de duas empresas criadas por ele e depois privatizadas que atendiam exclusivamente aos desejos do governo federal americano chamadas Fannie Mae e Freddie Mac, essas empresas possuíam vários privilégios concedidos pelo governo.

Junto com essas empresas, o governo criou uma espécie de cota para empréstimos, onde ele pressionava os bancos a emitirem empréstimos para pessoas que muitas vezes não podiam pagar, uma garantia de pagamento eram as hipotecas. Essas duas empresas compravam esses títulos hipotecários dos outros bancos e revendiam para bancos internacionais, isso fazia com

que cada vez mais os bancos americanos tivessem mais dinheiro para empréstimos, inflando ainda mais o mercado imobiliário.

No fim dos anos 90 e início dos anos 2000, com a bolha da internet e o atentado do 11 de setembro, havia um grande receio de uma possível crise e isso fez o governo abaixar a taxa de juros (Filgueiras, 2007), chegando em 1% ao ano em 2003, dessa forma os bancos aumentaram ainda mais o número de empréstimos e o preço dos imóveis disparou, e como os imóveis estavam sempre aumentando de preço as pessoas usavam o dinheiro da hipoteca para comprar carros, viajar, comprar roupas. Assim a economia dos EUA começou a depender do mercado imobiliário.

A partir de 2004, vendo que os sinais de crise haviam passado, o governo começou a aumentar os juros a níveis normais, em 2006 os preços dos imóveis estabilizaram e alguns até caíram, com esse aumento dos juros e diminuição dos preços dos imóveis as pessoas que tinham os imóveis hipotecados não pagavam mais e simplesmente deixavam para os bancos, entretanto o preço do imóvel não chegava nem perto do da hipoteca. Essa crise só se tornou global porque Fannie Mae e Freddie Mac vendiam os títulos hipotecários para bancos de todo o mundo, então agora aquelas pessoas que não pagaram as hipotecas estavam devendo para bancos do mundo inteiro, dessa forma gerando a maior crise desse século.

Em 2008 o então presidente Lula disse: “Lá (nos EUA), ela é um tsunami; aqui, se ela chegar, vai chegar uma marolinha que não dá nem para esquiar”, essa frase mostra que o governo vendia uma imagem que estávamos blindados e não sentiríamos as consequências aqui (Santos, 2012). Em geral os países emergentes não sentiram tanto a crise, entretanto apesar desse discurso a crise teve um impacto significativo no Brasil. Em outubro de 2008 a Sadia teve um prejuízo milionário por conta de títulos comprados por um valor maior que o real, isso levou a um prejuízo trimestral de mais de R\$ 2 bilhões de reais. Outra empresa que sofreu prejuízo foi a Aracruz com mais de R\$ 3 bilhões em perdas no último quarto de 2018.

Depois desses acontecimentos o governo tomou algumas medidas, houve diminuição na taxa básica de juros, SELIC, foi de 13,75% para 8,75% ao ano em 2009, dessa forma os juros pagos ao pegar um empréstimo eram menores e consequentemente circulava mais dinheiro. Diminuiu-se a alíquota de impostos (principalmente IPI) e liberou bilhões em depósitos compulsórios para os bancos, aumentando a liquidez (Pires, 2013). Mesmo que em 2008 o PIB nacional tenha aumentado 5,2%, em 2009 teve uma retração de 0,3%, a bolsa de valores, Bovespa, em 2008 teve uma queda de 4% a maior desde a década de 70.

Dessa forma através de modelos econométricos seria possível mensurar melhor o impacto dessa crise nos indicadores econômicos do Brasil e com isso entender o porquê dessas perdas enormes nas empresas e se essas medidas tomadas pelo governo surtiram efeito.

2 METODOLOGIA

A metodologia desenvolvida é baseada na área de modelagem econométrica, onde utilizamos análise de séries temporais para descrever o comportamento de indicadores econômicos durante o período da crise de 2008. Os dados mensais referentes ao período de janeiro de 2006 a dezembro de 2009 foram obtidos no site do Instituto de Pesquisa Aplicada (IPEA).

Foram escolhidos os seguintes indicadores para as análises:

- IPCA geral – índice (dez. 1993 = 100) – Índice de Preço ao Consumidor Amplo
- Índice de Preços das Exportações
- Índice de Preços das Importações
- Taxa de Juros Over/Selic
- Taxa de Câmbio Comercial

Foram realizadas análises descritivas para compreender melhor o comportamento dos indicadores e detectar possíveis tendências.

Após isso houve o ajuste de alguns modelos de decomposição de séries temporais utilizando dois tipos de modelos, um sendo um modelo de regressão de decomposição temporal e o outro de regressão harmônico, com isso foi feita a análise de resíduos para os modelos com maior potencial.

E por fim a partir da raiz quadrada do erro médio (RMSE), o erro médio absoluto (MAE), o ajuste do modelo e a análise dos resíduos, foi escolhido o melhor modelo.

Todas as análises foram feitas utilizando o software R, as etapas e as técnicas serão explicadas mais detalhadamente a seguir.

2.1 Modelo de Regressão Linear Simples

Uma série temporal é um conjunto de observações ordenadas no tempo, sendo assim é possível ajustar um modelo de regressão para uma série temporal.

Quando temos duas ou mais variáveis correlacionadas e o nosso interesse é entender os efeitos de uma sobre a outra, é possível através de uma expressão estimar os valores de uma das variáveis que chamamos de variável resposta ou de interesse, a partir dos valores das outras variáveis que chamamos de variáveis independentes ou explicativas. Esse método é denominado método de Regressão Linear e com ele podemos entender como as variáveis explicativas afetam a variável resposta.

Supondo uma variável resposta Y que sofre algum tipo de influência da variável explicativa X , podemos então ajustar um modelo de regressão simples

$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad (2.1)$$

onde β_0 é chamado de intercepto e β_1 é o coeficiente angular, ambos são constantes desconhecidas e ε um erro aleatório que deve atender alguns pressupostos, ele deve ser independente e identicamente distribuído com $E(\varepsilon) = 0$ e variância constante desconhecida σ^2 . Para a estimação de $\hat{\beta}_0$ e $\hat{\beta}_1$ é possível usar o método de mínimos quadrados, onde usamos a soma dos quadrados da diferença entre as observações y_i e o nosso modelo.

As estimativas de $\hat{\beta}_0$ e $\hat{\beta}_1$ serão dadas pelos pontos de máximos da função de mínimos quadrados

$$S(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2. \quad (2.2)$$

Desse modo chegamos aos seguintes estimadores de regressão

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \quad (2.3)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i)}{n}}{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}, \quad (2.4)$$

onde

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (2.5)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2.6)$$

Após essas estimativas devemos calcular também os resíduos ou erros do modelo que são dados por

$$\hat{\varepsilon}_i = y_i - \hat{y}_i. \quad (2.7)$$

Após calcularmos essas quantidades é possível construir uma tabela de análise de variância (ANOVA), onde é possível verificar a adequabilidade do modelo através de medidas como

- Soma de quadrado total (SQTotal), $\sum_{i=1}^n (y_i - \bar{y})^2$, que representa a distância quadrática entre os valores de y_i e a média $\bar{y}$.
- Soma de quadrados da regressão (SQReg), $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$, que representa a distância quadrática entre os valores ajustados $\hat{y}_i$ e a média $\bar{y}$.
- Soma de quadrado dos resíduos (SQRes), $\sum_{i=1}^n (y_i - \hat{y}_i)^2$, que representa a distância quadrática entre os valores de y_i e os valores estimados $\hat{y}_i$.

Além dessas medidas precisamos definir também os graus de liberdade para cada soma de quadrados. Os graus de liberdade são baseados no número de observações, de modo que a SQTotal possui $n - 1$ graus de liberdade, a SQReg possui 1 grau de liberdade e a SQRes possui $n - 2$.

Podemos também calcular o quadrado médio que é a razão entre a soma de quadrados e seus respectivos graus de liberdade, temos que o quadrado médio dos resíduos será nossa estimativa da variância $\hat{\sigma}^2$. Baseado nessas medidas temos que a tabela de análise de variância (ANOVA) será.

Tabela 1 – Análise de variância (ANOVA)

Fonte de variação	Graus de liberdade	Soma de quadrados	Quadrado Médio	F_0	P-valor
Regressão	1	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$QMReg = \frac{SQReg}{1}$	$\frac{QMReg}{QMRes}$	$P(F \geq F_0)$
Resíduo	$n - 2$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$QMRes = \frac{SQRes}{n-2}$		
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$			

Assim como mostrado na Tabela 1 precisamos calcular o p-valor, que é uma probabilidade do F tabelado ser maior ou igual ao observado, sendo que F_0 possui uma distribuição de probabilidade F de *Snedecor*, $F_0 \sim F_{(1,n-2)}$. Com isso podemos testar se o modelo é significativo com as seguintes hipóteses

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Rejeitamos a hipótese nula (H_0) se o p-valor for menor que um α determinado, normalmente utiliza-se $\alpha = 0.05$, ou seja se $P(F \geq F_0) \leq 0.05$ então podemos concluir que $\beta_1 \neq 0$, ou seja, o modelo é significativo e existe alguma tendência nos dados que pode ser explicada por uma reta.

Uma medida utilizada para checar a adequabilidade do modelo construído é o coeficiente de determinação (R^2), que é a razão entre a soma de quadrados da regressão e a soma de quadrados total, este coeficiente varia entre 0 e 1, sendo que quanto mais próximo de 1 mais o modelo está acertando.

2.2 Modelo de Regressão Linear Múltipla

No modelo de regressão múltipla iremos considerar novamente uma variável resposta Y porém dessa vez teremos duas ou mais variáveis explicativas, com isso nosso modelo para k variáveis explicativas será

$$Y_j = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \cdots + \beta_k X_{kj} + \varepsilon_j, \quad j = 1, 2, \dots, n$$

ou

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$
(2.8)

onde $\varepsilon \sim iid N(0, \sigma^2)$, a representação matricial será feita em negrito onde

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Os pressupostos são iguais aos já apresentados na regressão linear simples.

Para a estimação iremos utilizar novamente o método de mínimos quadrados, assim temos que as estimativas são dadas pelas equações, que tenta encontrar o vetor de estimadores que minimiza

$$Q = \sum_{j=1}^n \varepsilon_j^2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}),$$
(2.9)

Dessa forma temos as equações normais de mínimos quadrados

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y},$$
(2.10)

Para resolver as equações normais, multiplicamos ambos os lados pela inversa de $\mathbf{X}'\mathbf{X}$. Dessa forma temos que o estimador de mínimos quadrados de $\boldsymbol{\beta}$ é

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \quad (2.11)$$

Teremos então uma matriz coluna contendo as estimativas, a partir dessas estimativas e as matrizes construídas, assim como fizemos na regressão linear simples iremos montar a tabela de análise de variância para analisar a importância dos parâmetros para o modelo a partir de algumas medidas que serão dadas por

Soma de quadrados total (SQTotal) será:

$$\text{SQTotal} = \mathbf{y}'\mathbf{y} - \frac{(\sum_{j=1}^n y_j)^2}{n} \quad (2.12)$$

Soma de quadrados da regressão (SQReg) será:

$$\text{SQReg} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{(\sum_{j=1}^n y_j)^2}{n} \quad (2.13)$$

Soma de quadrados dos resíduos (SQRes) será:

$$\text{SQRes} = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \quad (2.14)$$

Nesse caso o quadrado médio do resíduo, que é a soma de quadrados dividida por seus graus de liberdades $(n - p)$, onde $p = k + 1$ parâmetros a serem estimados, também será a um estimador para a variância $\hat{\sigma}^2$.

Sabemos que adicionar uma variável no modelo provoca o aumento na soma de quadrados da regressão e a uma diminuição na soma de quadrados dos resíduos, assim devemos ter cuidado pois isso diminui a utilidade do modelo. Para testar a significância dos parâmetros de forma individual utilizamos o teste

$$\begin{aligned} H_0: \beta_j &= 0 \\ H_1: \beta_j &\neq 0 \quad j = 1, 2, \dots, n \end{aligned}$$

Se rejeitamos a hipótese nula então o parâmetro é significativo. O teste para essa hipótese é

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}}$$

em que C_{jj} é o elemento da diagonal de $(\mathbf{X}'\mathbf{X})^{-1}$ correspondendo a $\hat{\beta}_j$. Rejeitamos a hipótese nula H_0 se $|t_0|\sigma > t_{(\frac{\alpha}{2}, n-p)} \sim t \text{ de Student}(n-p)$.

Podemos ainda construir intervalos de confiança para os parâmetros, para isso vamos assumir que cada estatística $\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}}$ $j = 1, 2, \dots, n$, segue uma distribuição $t \text{ de Student}(n-p)$.

Baseado nisso podemos construir um intervalo de $100(1 - \alpha) \%$ de confiança para os coeficientes da regressão $\beta_j, j = 1, 2, \dots, n$ como

$$P\left(\hat{\beta}_j - t_{(\frac{\alpha}{2}, n-p)}\sqrt{\hat{\sigma}^2 C_{jj}} \leq \beta_j \leq \hat{\beta}_j + t_{(\frac{\alpha}{2}, n-p)}\sqrt{\hat{\sigma}^2 C_{jj}}\right) = 1 - \alpha \quad (2.15)$$

2.3 Decomposição de Modelos de Séries Temporais

Dessa vez iremos assumir que nossa variável resposta é uma série temporal Y_t , para $t = 1, 2, \dots, n$, onde ela varia no tempo, desse modo podemos escrever essa relação utilizando um modelo de regressão linear simples

$$Y_t = \beta_0 + \beta_1 t + a_t, \quad (2.16)$$

onde a_t é chamado *ruído branco* com média igual a zero, variância constante e sem covariância entre valores referentes a dois momentos distintos, ou seja,

$$E(a_t) = 0 \quad (2.17)$$

$$\begin{aligned} \text{Var}(a_t) &= \sigma_a \\ \gamma_k &= \text{cov}(a_t, a_{t+k}) = 0 \quad \text{para } k \neq 0 \end{aligned}$$

O modelo de decomposição de uma série temporal é dividido em 3 componentes: a tendência (D_t), a variação cíclica sazonal (S_t) e a parte aleatória (a_t)

$$Y_t = D_t + S_t + a_t \quad (2.18)$$

Dada uma série temporal Z_t podemos utilizar uma regressão linear para ajustar um modelo.

Para a tendência D_t umas das formas é utilizar uma forma polinomial

$$D_t = \beta_0 + \beta_1 t + \dots + \beta_m t^m, \quad (2.19)$$

onde o grau m do polinômio é bem menor que o número de observações N . Para estimar os parâmetros β_j utilizamos o método de mínimos quadrados, onde queremos minimizar a função

$$f(\beta_0, \dots, \beta_m) = \sum_{t=1}^N (Z_t - \beta_0 - \beta_1 t - \dots - \beta_m t^m)^2, \quad (2.20)$$

dessa forma obtemos os estimadores de mínimos quadrados usuais $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_m$.

Para a sazonalidade S_t os iremos usar a forma

$$S_t = \sum_{j=1}^{12} \alpha_j d_{jt}, \quad (2.21)$$

onde d_{jt} são variáveis periódicas, nesse caso iremos usar senos, cossenos e variáveis “dummy”.

Os métodos de estimação serão discutidos nos próximos tópicos

2.3.1 Teste de Dickey-Fuller

Uma série temporal pode ser ou não estacionária, ou seja, ela se mantém variando em torno de uma média ao longo do tempo. Isso é algo muito importante para a construção de modelos, e para verificar essa condição existe o teste de Dickey-Fuller (Moretin e Tolo, 2004) onde verificamos se a série é estacionária ou não, testando se possui raiz unitária.

Supondo uma série do tipo

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + a_t, \quad (2.22)$$

ao subtrairmos y_{t-1} de ambos os lados da igualdade, teremos que

$$\Delta y_t = \beta_0 + \beta_1 t + \lambda y_{t-1} + a_t \quad (2.23)$$

com $\lambda = \rho - 1$ e sendo a_t o ruído branco, com isso teremos que as hipóteses serão

$$H_0: \lambda = 0$$

$$H_1: -2 < \lambda < 0$$

Para esse teste temos três opções de aplicações, entretanto para o nosso objetivo iremos utilizar apenas uma delas.

Para a hipótese nula temos um modelo com tendência determinística linear, tendência estocástica e uma raiz unitária.

$$\Delta y_t = \beta_0 + \beta_1 t + a_t. \quad (2.24)$$

Para a hipótese alternativa, um processo sem essa raiz unitária

$$\Delta y_t = \beta_0 + \beta_1 t + \lambda y_{t-1} + a_t. \quad (2.25)$$

A estatística do teste é dada por

$$\tau = \frac{\hat{\lambda}}{s_{\hat{\lambda}}},$$

onde $s_{\hat{\lambda}}$ é o erro padrão de λ . Os valores críticos estão presentes em Mackinnon(1996).

2.3.1 Modelos de Decomposição com Variáveis Dummy

Uma variável “*dummy*” é aquela que só tem dois valores distintos, nesse caso utilizaremos -1 e 1. Para uma variável com k categorias teremos $k - 1$ variáveis “*dummy*”, com isso é possível distinguir dois períodos dentro da série temporal, nesse caso queremos distinguir os períodos pré e pós crise de 2008, desse modo nosso modelo é

$$Y_t = \beta_0 + \beta_1 t + \theta Z_{it} + \delta Z_{it} t + a_t, \quad t = 1, \dots, N; i = 1, \dots, t \quad (2.26)$$

sendo $Z_{it} = \begin{cases} -1 & \text{para o pré crise} \\ 1 & \text{para o pós crise} \end{cases}$

Olhando para a variável Z_{it} é possível concluir se houve diferença entre os dois períodos com relação ao intercepto, ou seja, o nível que a reta se localiza. Já se utilizarmos a variável $Z_{it}t$ verificamos se houve diferença com relação a inclinação da reta.

Iremos utilizar as variáveis “*dummy*” para cada ano, dada por

$$Z_{it} = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

E teremos a matriz $Z_{it}t$, para $t = 1, \dots, N$ e $i = 1, \dots, t$, que será a multiplicação de cada elemento da matriz Z_{it} por $t = 1, \dots, N$.

Para verificar se houve diferença entre os períodos com relação ao deslocamento da reta testamos

$$H_0: \theta = 0$$

$$H_1: \theta \neq 0,$$

já para verificar se houve diferença com relação a inclinação da reta testa-se

$$H_0: \delta = 0$$

$$H_1: \delta \neq 0.$$

Os testes de significância são feitos da mesma maneira apresentada na seção 2.2.

2.3.2 Modelos de Decomposição com Harmônicas

Utilizando funções periódicas (senos e cossenos) é possível modelar uma série que possua sazonalidade, essas harmônicas irão auxiliar no ajuste da curva por conta dessa periodicidade.

Iremos considerar o modelo de regressão matricial, como definido anteriormente na Equação 2.9, e $\mathbf{y}$ o vetor de observações. Suponha uma matriz de delineamento $\mathbf{X} = [\mathbf{T} : \mathbf{H}]$, tal que:

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t & t^2 & t^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & N & N^2 & N^3 \end{bmatrix},$$

$$\mathbf{H} = \begin{bmatrix} \cos(\omega_1) & \cdots & \cos(\omega_6) & \text{sen}(\omega_1) & \cdots & \text{sen}(\omega_6) \\ \cos(\omega_1 2) & \cdots & \cos(\omega_6 2) & \text{sen}(\omega_1 2) & \cdots & \text{sen}(\omega_6 2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(\omega_1 t) & \cdots & \cos(\omega_6 t) & \text{sen}(\omega_1 t) & \cdots & \text{sen}(\omega_6 t) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(\omega_1 N) & \cdots & \cos(\omega_6 N) & \text{sen}(\omega_1 N) & \cdots & \text{sen}(\omega_6 N) \end{bmatrix},$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}, \text{ e } \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \alpha_1 \\ \vdots \\ \alpha_6 \\ \gamma_1 \\ \vdots \\ \gamma_6 \end{bmatrix} = \begin{cases} \text{Intercepto} \\ \text{Tendência Linear} \\ \cos(\omega_1 t) \\ \vdots \\ \cos(\omega_6 t) \\ \text{sen}(\omega_1 t) \\ \vdots \\ \text{sen}(\omega_6 t) \end{cases},$$

$$\text{e } \omega_k = \frac{2\pi k}{12}, k = 1, \dots, 6 \text{ e } t = 1, \dots, N.$$

Foi utilizado para a tendência um polinômio de grau 3, por isso a matriz T possui 4 colunas, uma para cada parâmetro da tendência.

Sendo $\boldsymbol{\beta}$ o vetor dos parâmetros e $\boldsymbol{\varepsilon}$ o vetor de erros. Desse modo nosso modelo de regressão com harmônicas será:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \sum_{k=1}^6 \alpha_k \cos(\omega_k t) + \sum_{k=1}^6 \gamma_k \text{sen}(\omega_k t) + \varepsilon_t \quad (2.27)$$

As estimativas dos parâmetros são feitas pelo método de mínimos quadrados, dado por: $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, e as estimativas das observações $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$.

Para verificar a significância estatística do modelo, utilizamos os testes de hipóteses:

$$H_0: \beta_1 = \alpha_1 = \dots = \alpha_6 = \gamma_1 = \dots = \gamma_6 = 0$$

$$H_1: \text{pelo menos um parâmetro é diferente de zero}$$

Calcula-se a estatística F , que apresenta distribuição de *Fisher – Snedecor*, assim $F \sim F_{(p-1, n-p)}$. Rejeita-se H_0 ao nível de significância α se F observado for igual ou superior a F tabelado; ou se o p – valor for suficientemente pequeno, nesse caso, menor ou igual ao nível de significância $\alpha = 0,05$.

Para avaliar a significância dos parâmetros utilizaremos a mesma regra vista na Secção 2.2, onde testamos de forma individual cada parâmetro utilizando a estatística *t de Student*.

2.3.3 Modelos de Decomposição Temporal

Com o objetivo de contemplar a sazonalidade mensal, iremos considerar a mesma ideia de regressão utilizando harmônicos, porém para a decomposição temporal nosso modelo será:

$$Y = X\gamma + \varepsilon \quad (2.28)$$

onde Y é o vetor com as N=48 observações e $X = [C: D]$ é uma matriz de delineamento formada por:

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t & t^2 & t^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & N & N^2 & N^3 \end{bmatrix} \quad D = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix}, \text{ para}$$

$$D_{12 \times 11} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -1 & -1 & -1 & \dots & -1 \end{bmatrix} \begin{matrix} \leftarrow \text{jan} \\ \leftarrow \text{fev} \\ \vdots \\ \leftarrow \text{nov} \\ \leftarrow \text{dez} \end{matrix},$$

onde tem-se D com D_1, D_2, D_3 e D_4 por conta do período de 4 anos e para a tendência um polinômio de grau 3, por isso a matriz C possui 4 colunas, uma para cada parâmetro da tendência., temos também o vetor de parâmetros $\gamma = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$ e o vetor de erros $\varepsilon = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$. Nosso vetor de parâmetros possui a parte da tendência linear $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, com β_0 sendo a constante e β_1 o coeficiente angular e a sazonalidade mensal $\alpha' = [\alpha_1 \quad \dots \quad \alpha_{11}]$, dessa maneira o modelo será:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \sum_{i=1}^{11} \alpha_i d_{it} + \varepsilon_t \quad (2.29)$$

$$\text{para } d_t = \begin{cases} 1 & \text{para } i = \text{jan a dez} \\ -1 & \text{para } i = \text{dez} \\ 0 & \text{caso contrário} \end{cases}$$

A estimativa dos parâmetros é feita da mesma forma que o apresentado na secção 2.3.2, utilizando mínimos quadrados, assim $\hat{y} = (X'X)^{-1}X'y$ e o vetor de observações estimado é $\hat{Y} = X\hat{a}$.

Para avaliar a significância estatística do modelo, onde a hipótese nula H_0 estabelece que todos os parâmetros, incluindo β_1 e α_1 a α_{11} , são iguais a zero, enquanto a hipótese alternativa H_1 sugere que pelo menos um desses parâmetros é diferente de zero. Para conduzir esse teste, calcula-se a estatística F, que segue uma distribuição de Fisher-Snedecor sob a suposição da hipótese nula (ou seja, $F \sim F_{(p-1, n-p, 1-\alpha)}$)

A regra para rejeitar H_0 é quando a estatística F observada é maior ou igual ao valor crítico da distribuição F tabelado $F_{(p-1, n-p, 1-\alpha)}$, ou quando o pvalor associado ao teste é menor ou igual ao nível de significância escolhido, que, neste caso, é definido como $\alpha = 0,05$ para este estudo.

2.3.4 Mínimos Quadrados Generalizados

Nas secções anteriores utilizou-se a técnica de mínimos quadrados para o caso em que se tem homoscedasticidade e as covariâncias entre os erros eram iguais a zero, porém existem casos em que existe covariância positiva ou negativa entre os erros e para isso usa-se um caso mais geral (Hoffman, R. e Viera, 1987).

Para isso, consideramos o modelo

$$y = X\beta + u \quad (2.30)$$

com $E(u) = 0$ e $E(uu') = V\sigma^2$,

onde V é uma matriz $n \times n$ simétrica e definida positiva. Então V^{-1} também será simétrica e definida positiva, onde existe uma matriz Λ tal que

$$\Lambda'\Lambda = V^{-1} \quad (2.31)$$

onde Λ é dada por $\Lambda = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix},$

Multiplicando (2.1) por Λ obtemos o seguinte modelo

$$\Lambda y = \Lambda X \beta + \Lambda u \quad (2.32)$$

Agora podemos usar as mesmas expressões conhecidas para o modelo de regressão, porém agora de Λy contra ΛX , obtendo

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y, \quad (2.33)$$

que é um estimador linear não-tendencioso de variância mínima de β .

A matriz de variâncias e covariâncias de $\hat{\beta}$ é

$$(X'V^{-1}X)^{-1}\sigma^2 \quad (2.34)$$

A estimativa não tendenciosa de σ^2 é

$$s^2 = \frac{y'V^{-1}y - \hat{\beta}'X'V^{-1}y}{n - p} \quad (2.35)$$

A partir dessas estimativas as expressões do modelo de regressão múltipla podem ser generalizadas apenas substituindo X por ΛX e y por Λy e lembrando que $\Lambda'\Lambda = V^{-1}$.

2.4 Análise de Resíduos

A partir do momento em que construímos nosso modelo, devemos checar os resíduos, que por definição são as n diferenças entre os valores reais e os estimados pelo modelo, $e_i = Y_i - \hat{Y}_i$, $i = 1, 2, \dots, n$, ou seja, a quantidade que o modelo não foi capaz de explicar.

Os resíduos devem atender alguns pressupostos importantes, segundo Montgomery, Peck e Vining (2006) são eles:

1. A relação entre a variável resposta y e as explicativas deve ser aproximadamente linear.
2. Os resíduos $\hat{\varepsilon}$ devem ter média zero, $E(\hat{\varepsilon}) = 0$.
3. Os resíduos $\hat{\varepsilon}$ devem ter variância σ^2 constante.
4. Os resíduos devem ser não correlacionados
5. Os resíduos devem ser normalmente distribuídos, $\hat{\varepsilon} \sim N(0, \sigma^2)$.

A partir desses pressupostos podemos verificar a adequabilidade do nosso modelo, para isso utilizaremos alguns testes.

2.4.1 Teste de Shapiro-Wilk

Para verificarmos a normalidade dos resíduos é possível realizar o teste Shapiro-Wilk, onde suas hipóteses são:

H_0 : os resíduos seguem distribuição normal

H_1 : os resíduos não seguem distribuição normal

E para o teste iremos usar a estatística dada por:

$$W_{obs} = \frac{[\sum_{i=1}^k a_{n-i+1}(y_{n-i+1} - y_i)]^2}{[\sum_{i=1}^n (y_i - \bar{y})^2]^2} \quad (2.36)$$

Para a interpretação dos resultados é necessário a utilização de tabelas que estão presentes em Shapiro e Wilk (1965).

Rejeita-se H_0 se $W_{tab} < W_{obs}$, ao nível de significância α .

2.4.2 Teste de Durbin-Watson

Outra suposição que fazemos é a de independência entre os resíduos, para verificar isso iremos aplicar o teste de Durbin-Watson que é mostrado em Draper e Smith (1968).

Vamos supor um modelo dado por:

$$y_u = \beta_0 + \sum_{i=1}^k \beta_i X_{iu} + \varepsilon_u, \quad u = 1, 2, \dots, n \quad (2.37)$$

testaremos as seguintes hipóteses:

$$H_0: \rho_s = 0$$

$$H_1: \rho_s \neq 0$$

onde ρ_s é a correlação serial dos resíduos. A estatística para o teste é:

$$d = \frac{\sum_{u=2}^n (\varepsilon_u - \varepsilon_{u-1})^2}{\sum_{u=1}^n \varepsilon_u^2} \quad (2.38)$$

Entretanto a para rejeitar ou não a hipótese nula, é necessário olhar para dois valores críticos, d_L e d_U que estão presentes em Draper e Smith (1968), e a partir disso usaremos a seguinte regra de decisão para um teste bicaudal com nível de significância α .

Se $d < d_L$ ou $4 - d < d_L$, rejeitamos a hipótese nula H_0 ao nível de significância 2α .

Se $d > d_U$ e $4 - d > d_U$, não rejeitamos a hipótese nula H_0 ao nível de significância 2α .

Caso contrário, o teste é inconclusivo.

2.4.3 Análises Gráficas

Além dos testes mostrados anteriormente, podemos verificar alguns pressupostos dos resíduos baseando-se na análise gráfica, que é uma forma mais simples de verificar se os pressupostos estão sendo violados ou não. Alguns gráficos de resíduos úteis são, resíduos versus valores preditos e resíduos em ordem temporal (se for conhecida), além do gráfico de dispersão básico.

Ao fazermos esses gráficos esperamos que os resíduos tenham um certo comportamento para que possamos concluir que nenhum pressuposto foi violado. Esperamos que os resíduos estejam distribuídos aleatoriamente ao longo de uma “reta” horizontal, isso se dá para todos os gráficos. A partir dessa distribuição dos resíduos é possível checar se nosso modelo está explicando bem os dados, caso contrário nossos resíduos poderão ter características que irão nos ajudar a ajustar o modelo.

2.5 Medidas para o Diagnóstico dos Modelos

Para a avaliação do desempenho dos modelos foram utilizadas as medidas do Erro Médio Absoluto (MAE) e a raiz quadrada do Erro Quadrático Médio (RMSE), que são obtidos a partir da estimativa $\hat{y}_i$ e dos valores observados y_i , e são calculadas da seguinte forma:

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{N}}, \quad (2.39)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (2.40)$$

Quanto mais próximo de zero as medidas, melhor a previsão do modelo.

A medida MAE é mais sensível a valores discrepantes se comparada ao RMSE, isso se deve pelo uso do módulo, ou seja, ele dá pesos iguais a todos os desvios.

O RMSE dá um peso maior a desvios grandes já que estão elevados ao quadrado.

3 RESULTADOS

A partir das técnicas descritas anteriormente e dos dados coletados do IPEA, foram construídos dois modelos de regressão temporal dos indicadores econômicos para avaliar seu comportamento durante o período 2006-2009 e com isso avaliar se a crise financeira de 2008 teve um impacto significativo na tendência das séries. As análises foram feitas utilizando o software R, e serão apresentadas na seção seguinte.

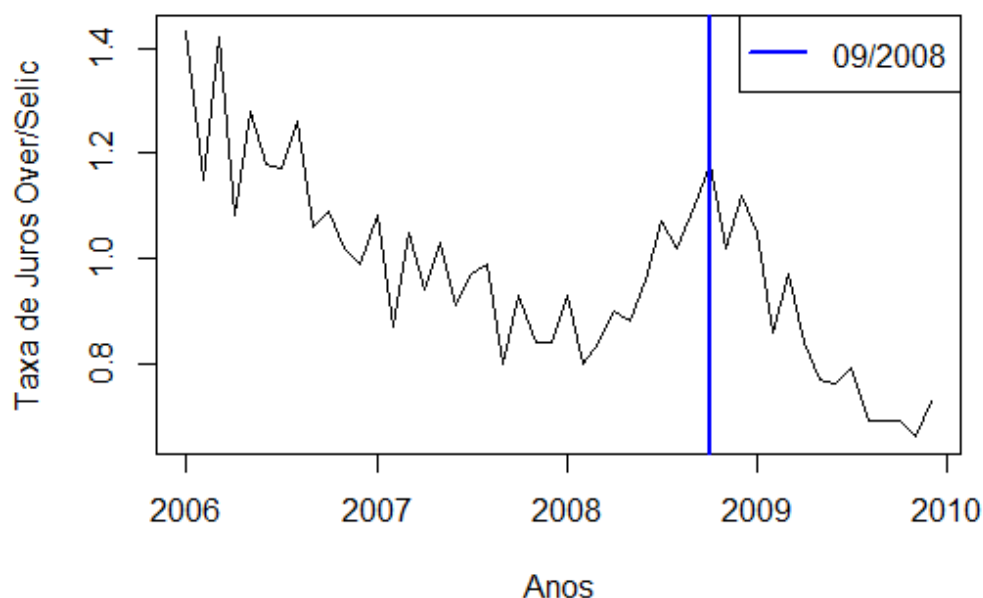
3.1 Análises Descritivas dos Indicadores

Para a análise descritiva iremos utilizamos medidas como média, desvio padrão e o coeficiente de variação, além de visualização gráfica da série para verificar se existe alguma tendência ou sazonalidade e avaliar seu comportamento.

Na figura 1, vemos que a Taxa de Juros já vinha em queda desde 2006, a partir de 2008 houve uma crescente e com a quebra do banco Lehman Brothers em 09/2008 voltou a cair novamente.

Analisando a série vemos que aparentemente possui uma tendência de queda, ou seja, não é estacionária. A partir do teste de Dickey-Fuller com um nível de significância de $\alpha = 0,05$, chegamos em um p-valor de 0,5428, concluindo assim que a série realmente não é estacionária.

Figura 1 – Série da Taxa de Juros Over/Selic



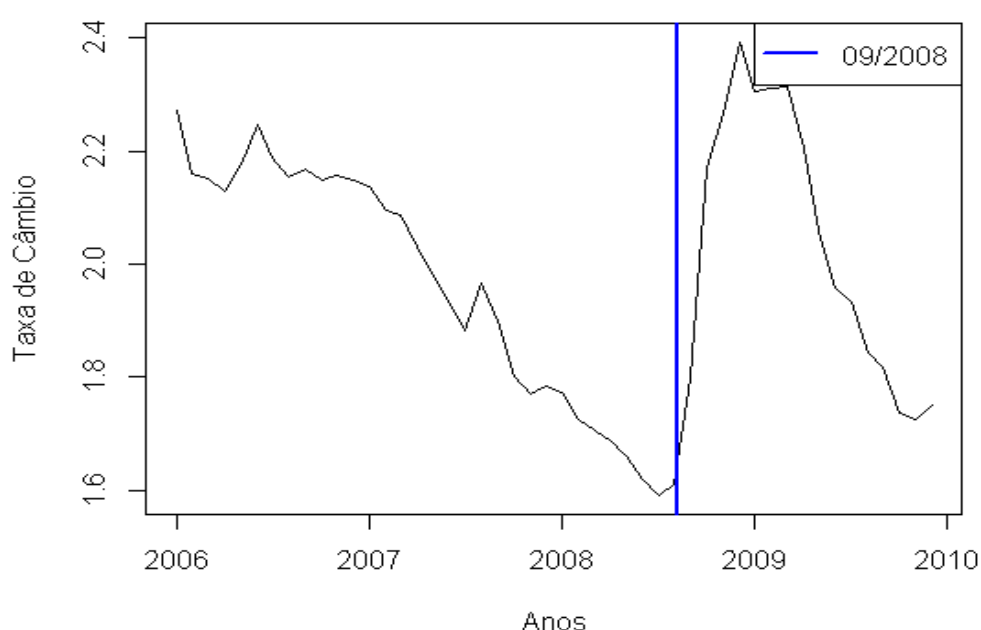
Pela tabela 1, verificamos de forma algumas medidas relacionadas à série, pelas médias ao longo dos anos vemos que realmente houve uma queda da Taxa de Juros, de 2007 a 2008 houve uma pequena crescente, entretanto, a partir do estopim da crise em novembro de 2008 a média cai bastante. Vemos que o desvio padrão cai drasticamente em 2007, entretanto sobe novamente nos anos seguintes, e de certa forma se mantém constante. A partir do coeficiente de variação ($\frac{\text{desvio padrão}}{\text{média}}$) vemos que não há uma grande variação dos dados em relação à média.

Tabela 1 - Estatísticas da Taxa de Juros

	2006	2007	2008	2009
<i>Média</i>	1,177	0,9375	0,985	0,7917
<i>Desvio Padrão</i>	0,145	0,090	0,119	0,120
<i>CV</i>	12,32%	9,20%	12,10%	15,19%

Na figura 2, percebe-se que a Taxa de Câmbio vinha caindo desde 2006 e um pouco antes de setembro de 2008 começa a subir rapidamente possivelmente por conta da crise que teve seu estopim em 09/2008, até que no início de 2009 volta a cair. Realizando o teste de Dickey-Fuller com um nível de significância de $\alpha = 0,05$, chega-se a um p-valor de 0,4564 e conclui-se que a série não é estacionária.

Figura 2 – Série da Taxa de Câmbio



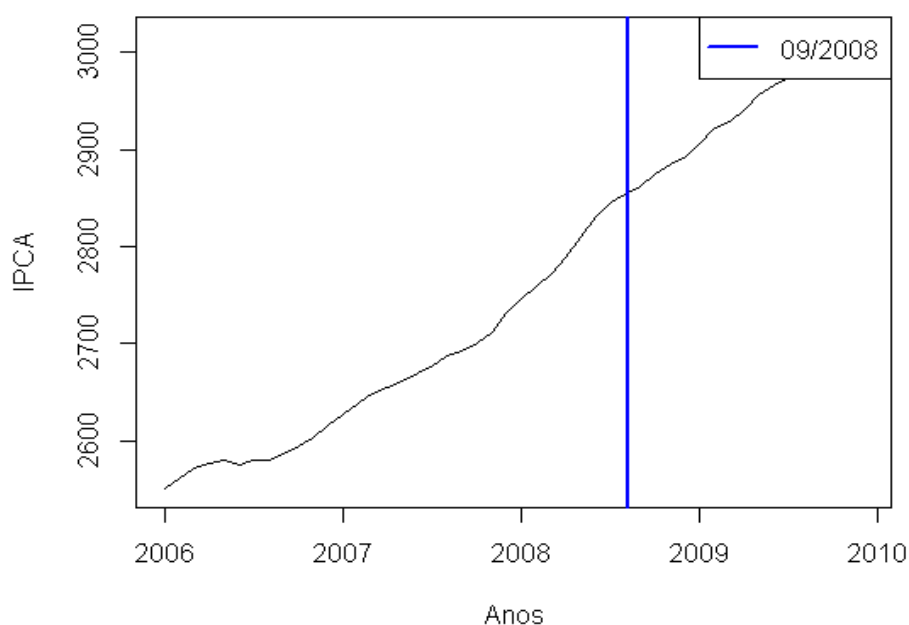
Pela tabela 2, as médias se mantiveram bem próximas ao longo dos quatro anos, ao olhar para o coeficiente de variação percebe-se um aumento seguido nos anos de 2006, 2007 e 2008 e uma queda em 2009, porém nos anos de 2006 e 2007 este coeficiente é baixo (menores que 10%) e nos anos de 2008 e 2009 pode ser considerado médio (entre 10% e 20%), o que indica que os dados não estão tão dispersos em relação à média nos quatro anos.

Tabela 2 - Estatísticas da Taxa de Câmbio

	2006	2007	2008	2009
<i>Média</i>	2,175	1,947	1,892	1,997
<i>Desvio Padrão</i>	0,043	0,125	0,278	0,235
<i>CV</i>	1,97%	5,92%	15,19%	11,75%

Na figura 3, nota-se que o IPCA se manteve em uma tendência crescente e aparentemente não sofreu alteração por conta da crise, pelo teste de Dickey-Fuller obteve-se um p-valor de 0,09345, o que indica que a série não é estacionária.

Figura 3 – Série do IPCA



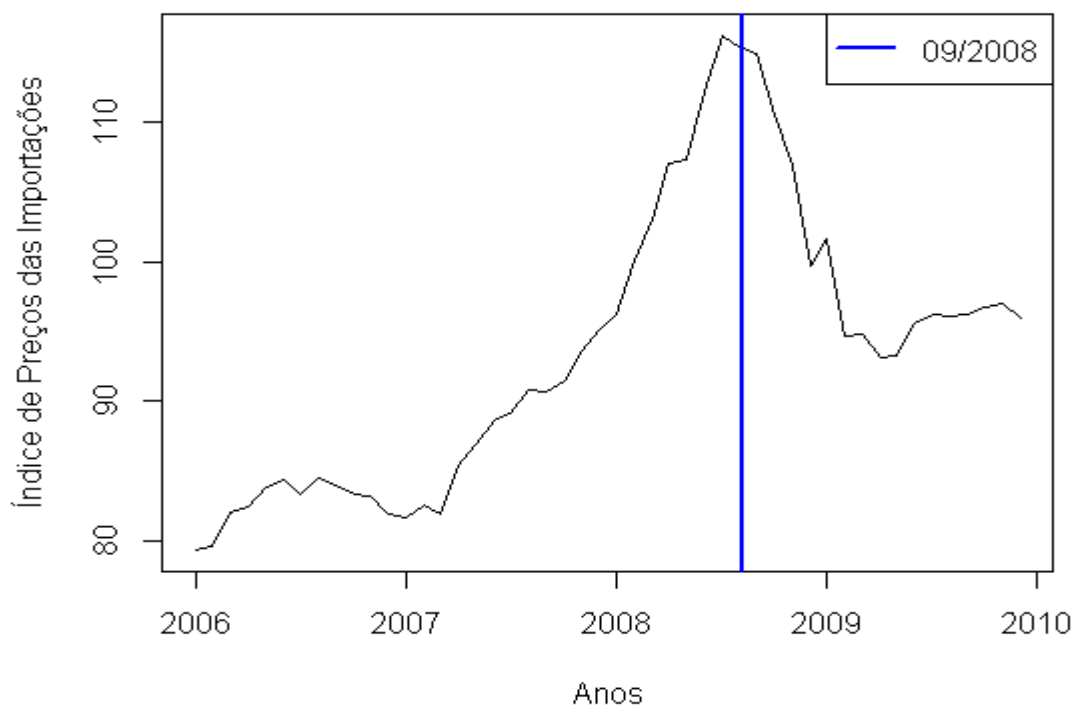
Na tabela 3, a média aumentou em todos os anos e ao analisar o coeficiente de variação nota-se que todos são bem baixos (menores que 10%) o que indica que os dados não estão dispersos e sim em torno da média, ou seja, a média é uma medida significativa.

Tabela 3 - Estatísticas do IPCA

	2006	2007	2008	2009
<i>Média</i>	2581	2675	2827	2965
<i>Desvio Padrão</i>	17,385	31,346	50,391	34,548
<i>CV</i>	0,67%	1,20%	1,78%	1,16%

Na figura 4, o Índice de Preços das Importações vinha subindo desde 2006 e a partir de setembro de 2008 começa a cair rapidamente possivelmente por conta da crise que fez com que o dólar desvalorizasse e os preços das importações diminuíssem. Realizando o teste de Dickey-Fuller obteve-se um p-valor de 0,4061, logo a série não é estacionária.

Figura 4 – Série do Índice de Preços das Importações



Na tabela 4 é possível notar que a média teve um aumento em 2008, porém em 2009 já houve uma queda por conta da crise que fez os preços caírem com a desvalorização do dólar,

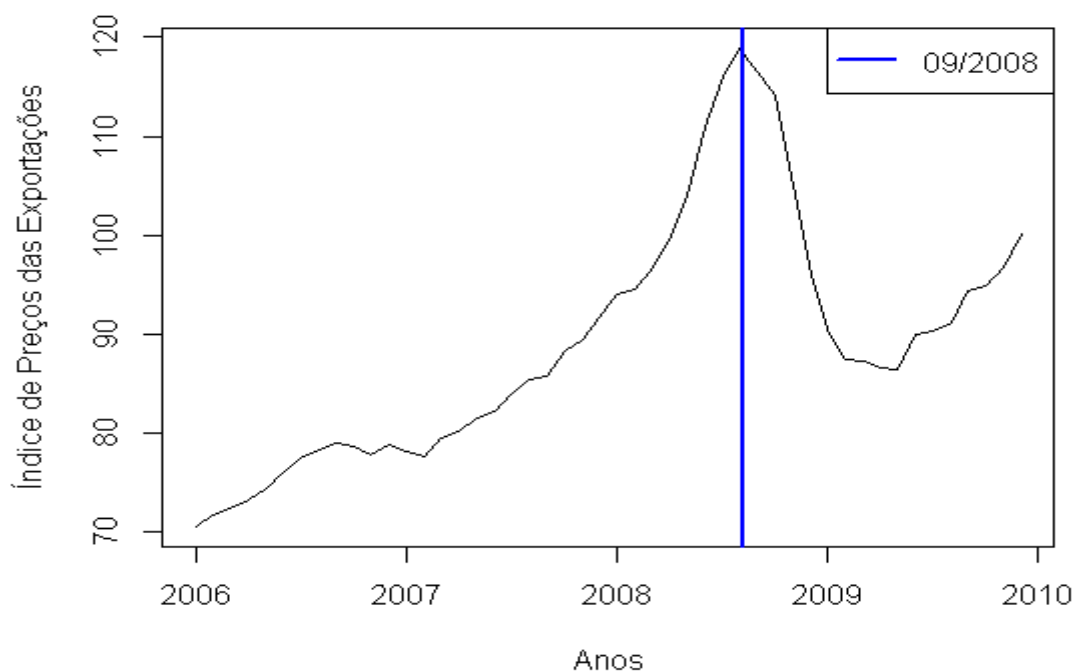
ao analisar o coeficiente de variação nota-se que todos são bem baixos (menores que 10%) o que indica que os dados não estão dispersos e sim em torno da média.

Tabela 4 - Estatísticas do Índice de Preços das Importações

	2006	2007	2008	2009
<i>Média</i>	82,66	88,16	107,48	95,95
<i>Desvio Padrão</i>	1,703	4,534	6,662	2,163
<i>CV</i>	2,06%	5,52%	6,20%	2,25%

Na figura 5, o Índice de Preços das Exportações vinha crescendo até 2008 e com a crise houve uma queda por conta da desvalorização do dólar e a desaceleração da economia dos países desenvolvidos como os EUA e países da Europa. Pelo teste de Dickey-Fuller obteve-se um p-valor de 0,3221 o que indica que a série não é estacionária.

Figura 5 – Série do Índice de Preços das Exportações



Na tabela 5 é possível notar que a média teve um aumento em 2008, porém em 2009 já houve uma queda por conta da crise que fez o dólar desvalorizar e a desaceleração da economia dos países como os EUA e os países da Europa, ao analisar o coeficiente de variação nota-se

que todos são bem baixos (menores que 10%) o que indica que os dados não estão dispersos e sim em torno da média.

Tabela 5 - Estatísticas do Índice de Preços das Exportações

	2006	2007	2008	2009
<i>Média</i>	75,65	83,61	105,54	91,28
<i>Desvio Padrão</i>	3,101	4,585	9,421	2,162
<i>CV</i>	4,10%	5,90%	8,93%	4,80%

3.2 Modelagem e Análise dos Indicadores

Para a construção dos modelos foram utilizadas variáveis “dummy” Z_{it} , onde assumimos os valores -1 para o período pré-crise, de Janeiro de 2006 até Setembro de 2008, e 1 para o período pós crise, de Outubro de 2008 até Dezembro de 2009, para a sazonalidade utilizou-se harmônicos e uma decomposição temporal.

Os parâmetros foram estimados pelos métodos de mínimos quadrados ordinários e para aqueles que apresentaram autocorrelação residual aplicou-se mínimos quadrados generalizados e foram utilizados apenas aqueles que foram significativos a um nível de significância de $\alpha = 0,05$.

Será utilizado Modelo 1 para referir-se ao modelo utilizando a matriz $X = [T : H]$, com harmônicos apresentado na seção 2.3.2 e Modelo 2 para o modelo utilizando a matriz $X = [C : D]$, mostrado na seção 2.3.3

3.2.1 Modelagem da Taxa de Juros Over/Selic

As estimativas de mínimos quadrados generalizados com $\hat{\rho} = 0,75$ desses parâmetros para o Modelo 1 usando harmônicos da série de Taxa de Juros Over/Selic estão na Tabela 6.

Tabela 6 – Estimação dos Parâmetros do Modelo 1 da Taxa de Juros

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	3,5385	$9,7333 \times 10^{-1}$	$1,3327 \times 10^{-3}$
β_1	-2,992	$3,0792 \times 10^{-2}$	$5,499 \times 10^{-3}$
α_6	-2,3169	$1,4831 \times 10^{-2}$	$2,7518 \times 10^{-2}$
γ_4	2,2194	$1,15 \times 10^{-2}$	$3,4163 \times 10^{-2}$
γ_5	4,9904	$1,0583 \times 10^{-2}$	$2,3937 \times 10^{-5}$
θ	2,0962	$9,8487 \times 10^{-1}$	$4,4594 \times 10^{-2}$
δ	-2,0887	$2,9175 \times 10^{-2}$	$4,5316 \times 10^{-2}$

Baseado nas estimativas e nos parâmetros significativos α_6 , γ_4 e γ_5 podemos concluir que existe sazonalidade na série da Taxa de Juros, vemos que a estimativa de β_1 (coeficiente angular) é negativa, ou seja, a série tem uma tendência negativa. Temos também que os parâmetros θ e δ foram significativos no Modelo 1, ou seja, existe uma diferença tanto no intercepto quanto na inclinação na Taxa de Juros entre os períodos pré e pós crise.

O Modelo 1 da Taxa de Juros final é dado por:

$$\hat{y}_t = 3,5385 - 2,992t + 2,2194 \cos\left(\frac{12\pi t}{12}\right) + 2,2194 \sin\left(\frac{8\pi t}{12}\right) + 4,9904 \sin\left(\frac{10\pi t}{12}\right) + 2,0962Z_{it} - 2,0887Z_{it}t$$

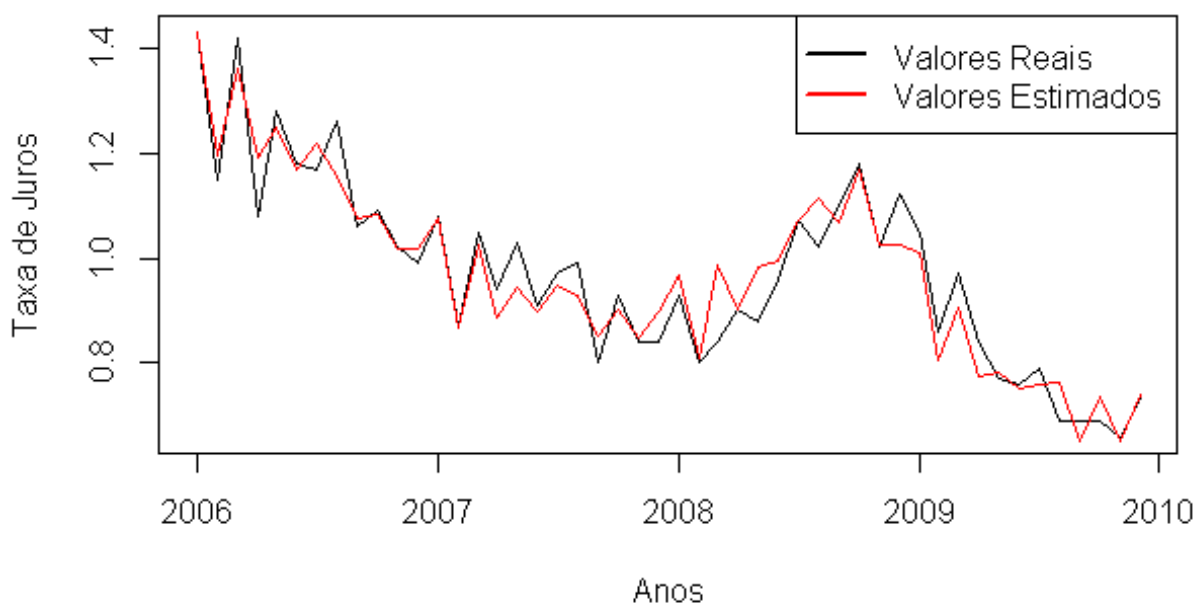
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 1,4423 - 0,9033t + 2,2194 \cos\left(\frac{12\pi t}{12}\right) + 2,2194 \sin\left(\frac{8\pi t}{12}\right) + 4,9904 \sin\left(\frac{10\pi t}{12}\right)$$

Já para o pós-crise teremos

$$\hat{y}_t = 5,6347 - 5,0807t + 2,2194 \cos\left(\frac{12\pi t}{12}\right) + 2,2194 \sin\left(\frac{8\pi t}{12}\right) + 4,9904 \sin\left(\frac{10\pi t}{12}\right)$$

Figura 6 – Modelo 1 da Taxa de Juros Ajustado



Vemos que o Modelo 1 está se ajustando bem a série, mesmo após a queda brusca em 2008 por conta da ação do governo para tentar aumentar a liquidez do mercado, injetando capital no mercado.

Para o Modelo 2 usando decomposição temporal também foi utilizado o método de mínimos quadrados generalizados com $\hat{\rho} = 0,8$ para a estimação dos parâmetros e foram utilizados apenas aqueles que foram significativos a um nível de significância de $\alpha = 0,05$.

As estimativas desses parâmetros para o Modelo 2 da série de Taxa de Juros Over/Selic estão na Tabela 7.

Tabela 7 – Estimação dos Parâmetros do Modelo 2 da Taxa de Juros

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	3,4427	$8,6517 \times 10^{-1}$	$3,71 \times 10^{-4}$
β_1	$-9,5839 \times 10^{-2}$	$3,394 \times 10^{-2}$	$8,1 \times 10^{-3}$
β_3	$2,471 \times 10^{-5}$	$1,0297 \times 10^{-5}$	$2,24 \times 10^{-2}$
α_2	$-1,094 \times 10^{-1}$	$4,6291 \times 10^{-2}$	$2,42 \times 10^{-2}$
θ	2,0645	$9,8487 \times 10^{-1}$	$1,628 \times 10^{-2}$
δ	$-6,094 \times 10^{-2}$	$2,9175 \times 10^{-2}$	$1,742 \times 10^{-2}$

Baseado nas estimativas e no parâmetro significativo α_2 podemos concluir que o modelo identificou sazonalidade na série da Taxa de Juros, a estimativa de β_1 é negativa, ou seja, a série tem uma tendência negativa. Os parâmetros θ e δ também foram significativos o que indica que existe uma diferença tanto no intercepto quanto na inclinação. Ambos os modelos para a Taxa de Juros possuem estimativas dos parâmetros bem próximas, e chegaram as mesmas conclusões a respeito da série.

O Modelo da Taxa de Juros final para o indicador é dado por:

$$\hat{y}_t = 3,4427 - 9,5839 \times 10^{-2}t + 2,471 \times 10^{-5}t^3 - 1,094 \times 10^{-1}d_{1t} + 2,0645Z_{it} - 6,094 \times 10^{-2}Z_{it}t$$

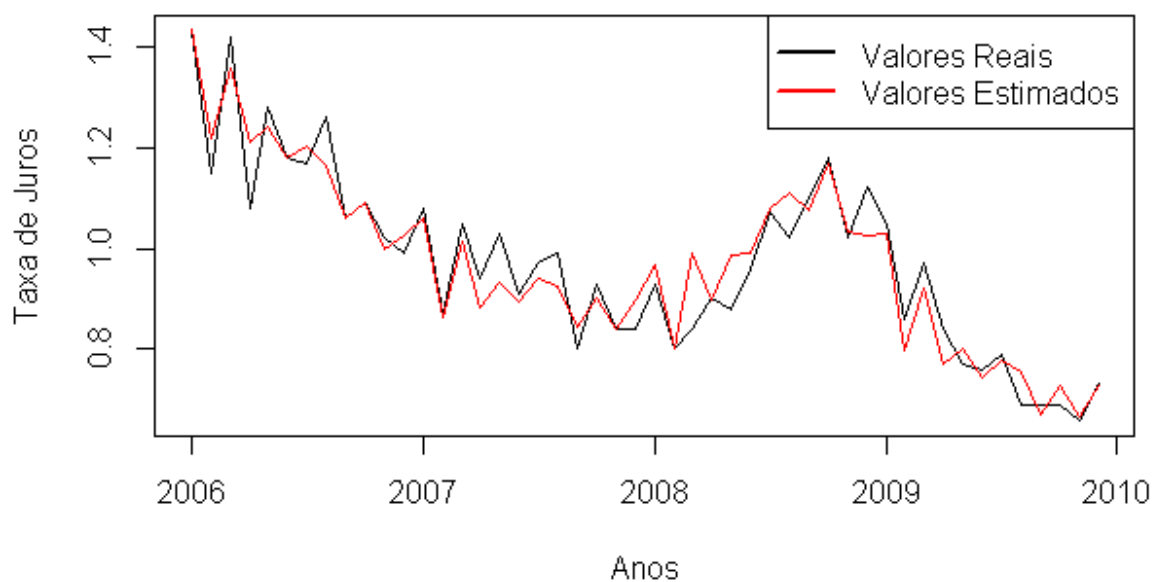
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 2,06457 - 3,4899 \times 10^{-2}t + 2,471 \times 10^{-5}t^3 - 1,094 \times 10^{-1}d_{1t}$$

Já para o pós-crise teremos

$$\hat{y}_t = 5,55072 - 15,6779 \times 10^{-2}t + 2,471 \times 10^{-5}t^3 - 1,094 \times 10^{-1}d_{1t}$$

Figura 7 – Modelo 2 da Taxa de Juros Ajustado



O ajuste está bem semelhante ao do modelo utilizando harmônicos, ambos estão se ajustando muito bem a série e estão conseguindo contemplar muito bem a tendência e a sazonalidade, mesmo com a variação de 2008 e 2009.

3.2.2 Análise dos Resíduos e Adequabilidade dos Modelos da Taxa de Juros Over/Selic

As figuras 8 e 9 representam os resíduos dos modelos da Taxa de Juros

Figura 8 – Resíduos do Modelo 1

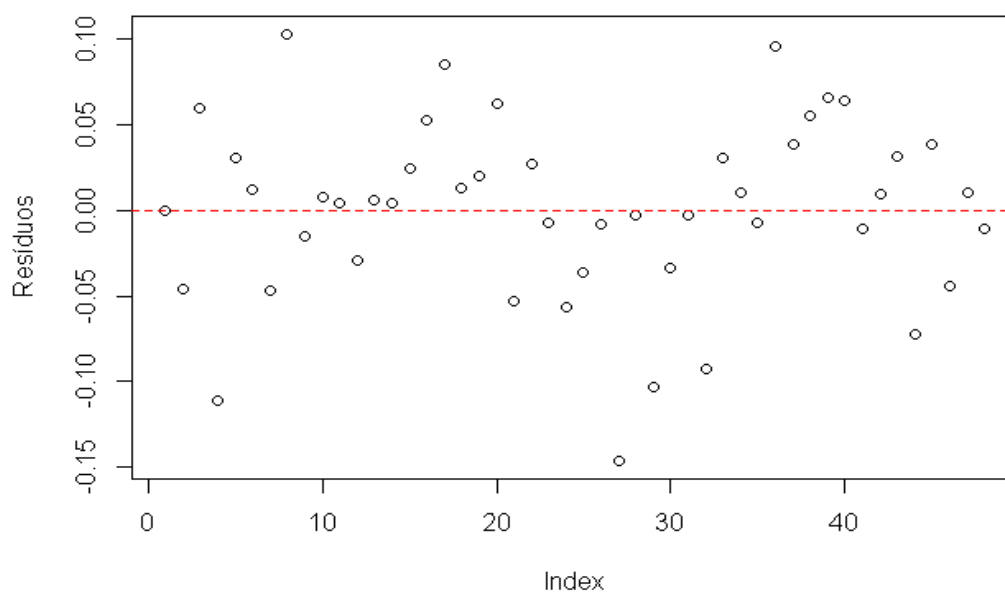
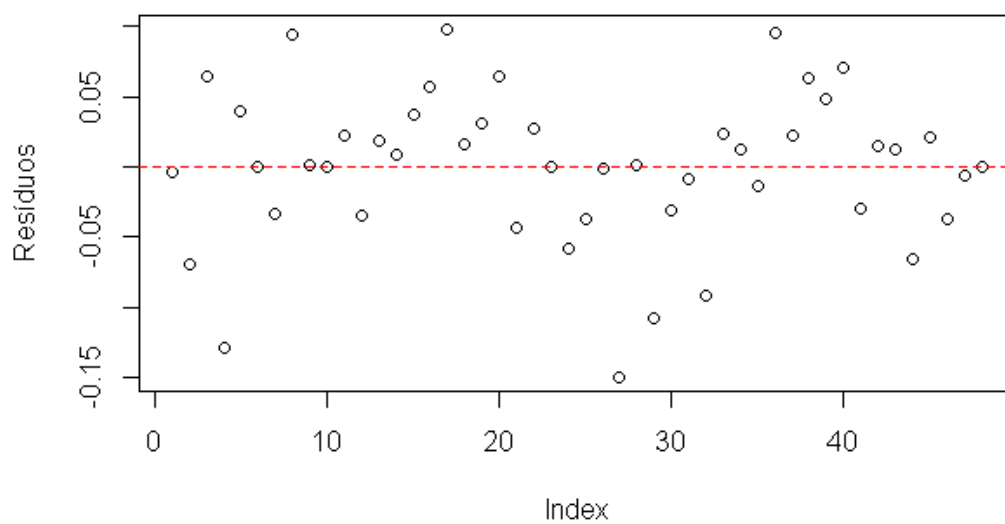


Figura 9 – Resíduos do Modelo 2



Ambos os gráficos aparentemente apresentam uma certa tendência e não parecem se comportarem de forma aleatória em torno do zero (linha tracejada). Em ambos os modelos se nota uma pequena ondulação em torno da linha tracejada o que pode indicar uma certa autocorrelação, para isso utilizaremos o teste de Durbin-Watson que está representado na Tabela 8.

Tabela 8 – Teste de Durbin-Watson

	<i>Estatística</i>	<i>dl</i>	<i>du</i>
<i>Modelo 1</i>	$DW = 2,076$	1,246	1,875
<i>Modelo 2</i>	$DW = 2,08$	1,291	1,822

Temos na tabela a estatística do teste e os limites inferiores e superiores tabelados para um $\alpha = 0,05$ para realizar o teste. Como no Modelo 1 $DW > du$ então não se rejeita a hipótese nula de que $\rho = 0$, e conclui-se que não existe autocorrelação. Para o Modelo 2 temos que $DW > du$, dessa forma não se rejeita a hipótese nula e concluímos que também não existe autocorrelação

É necessário também verificarmos a normalidade dos resíduos, a seguir temos a tabela com a estatística e o p-valor do teste de Shapiro-Wilk.

Tabela 9 – Teste de Shapiro-Wilk

	<i>Estatística</i>	<i>p – valor</i>
<i>Modelo 1</i>	$W = 0,973$	0,3323
<i>Modelo 2</i>	$W = 0,9623$	0,125

Pela Tabela 9 vemos que os resíduos apresentaram $p - valor > 0,05$, dessa forma não rejeitamos a hipótese nula a um nível de significância $\alpha = 0,05$ e concluímos que os resíduos seguem uma distribuição normal.

Com isso podemos assumir que nosso modelo está adequado, agora precisamos verificar a acurácia destes modelos.

Para isso vamos usar medidas como o RMSE, o MAE e o R^2 .

Tabela 10 – Medidas de Diagnóstico

	<i>RMSE</i>	<i>MAE</i>	R^2
<i>Modelo 1</i>	0,0524	0,1990	0,9148
<i>Modelo 2</i>	0,0538	0,20	0,9102

A partir da Tabela 10 vemos que os modelos estão performando muito bem, ambas medidas para verificar o erro cometido estão baixas e bem próximas, entretanto o Modelo 1 possui ambas medidas menores, o coeficiente de determinação que mede o quanto o modelo está explicando a série também é maior no Modelo 1, com 91,48% da variabilidade explicada

3.2.3 Modelagem da Taxa de Câmbio

A seguir temos as estimativas de mínimos quadrados generalizados dos parâmetros com $\hat{\rho} = -0,32$ para o Modelo 1 com harmônicos da série da Taxa de Câmbio na Tabela 11.

Tabela 11 – Estimação dos Parâmetros do Modelo 1 da Taxa de Câmbio

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	10,0965	$3,3522 \times 10^{-1}$	$3,6564 \times 10^{-11}$
β_2	-2,3876	$7,3922 \times 10^{-4}$	$2,346 \times 10^{-2}$
θ	3,39641	$3,607 \times 10^{-1}$	$1,9418 \times 10^{-3}$
δ	-2,1124	$1,0594 \times 10^{-2}$	$4,202 \times 10^{-2}$

Para esta série vemos que nenhum harmônico foi significativo, apenas o intercepto β_0 e β_2 que indica uma tendência quadrática negativa, além disso ambos os parâmetros das variáveis “dummy” foram significativos, o que mostra que a crise afetou tanto a inclinação da série com a estimativa de θ negativa o que mostra uma queda na tendência da Taxa de Câmbio, quanto no intercepto indicado por δ .

O Modelo 1 da Taxa de Câmbio final é dado por:

$$\hat{y}_t = 10,0965 - 2,3876t^2 + 3,9641Z_{it} - 2,1240Z_{it}t$$

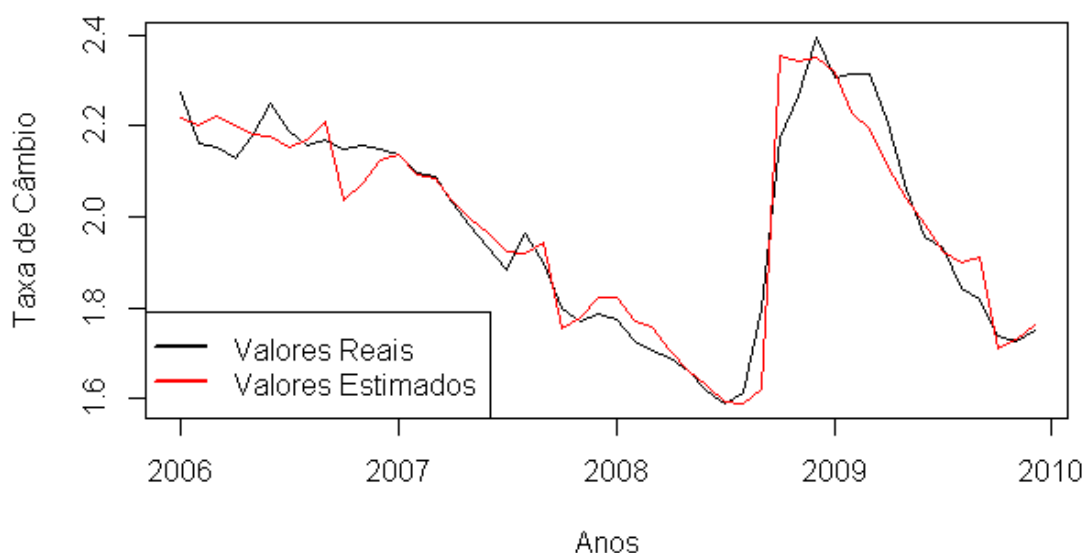
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 6.1324 + 2,1240t - 2,3876t^2$$

Já para o pós-crise teremos

$$\hat{y}_t = 14.0606 - 2,1240t - 2,3876t^2$$

Figura 10 – Modelo 1 da Taxa de Câmbio Ajustado



Nota-se um bom ajuste, no período antes da crise percebe-se uma tendência negativa, ou seja, o real vinha caindo de valor em relação ao dólar, já após a crise vemos um grande salto por conta da desvalorização da moeda americana e conseqüentemente uma valorização do real

em relação ao dólar. Entretanto logo após medidas do governo para abaixar as taxas de juros injetando dinheiro na economia, a inflação aumentou e dessa forma o real voltou a se desvalorizar em comparação ao dólar.

Para o Modelo 2 usando decomposição temporal também foi utilizado o método de mínimos quadrados generalizados a um nível de significância de $\alpha = 0,05$.

As estimativas desses parâmetros para o Modelo 2 com $\hat{\rho} = -0,8$ da série de Taxa de Câmbio estão na Tabela 12.

Tabela 12 – Estimação dos Parâmetros do Modelo 2 da Taxa de Câmbio

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	11,1804	$3,0704 \times 10^{-1}$	$2,09 \times 10^{-12}$
β_2	-2,8063	$7,8198 \times 10^{-3}$	$8,58 \times 10^{-3}$
β_3	2,1611	$1,2682 \times 10^{-5}$	$3,85 \times 10^{-2}$
θ	3,8773	$3,3215 \times 10^{-1}$	$5,13 \times 10^{-4}$
δ	-2,4706	$9,7598 \times 10^{-3}$	$1,92 \times 10^{-2}$

Assim como no Modelo 1 vemos que a série não possui a componente sazonal apenas a tendência, temos que β_2 também é negativo, porém neste caso foi necessário o parâmetro β_3 para um bom ajuste.

O Modelo 2 da Taxa de Câmbio final é dado por:

$$\hat{y}_t = 11,1804 - 2,8063t^2 + 2,1611t^3 + 3,8773Z_{it} - 2,4706Z_{it}t$$

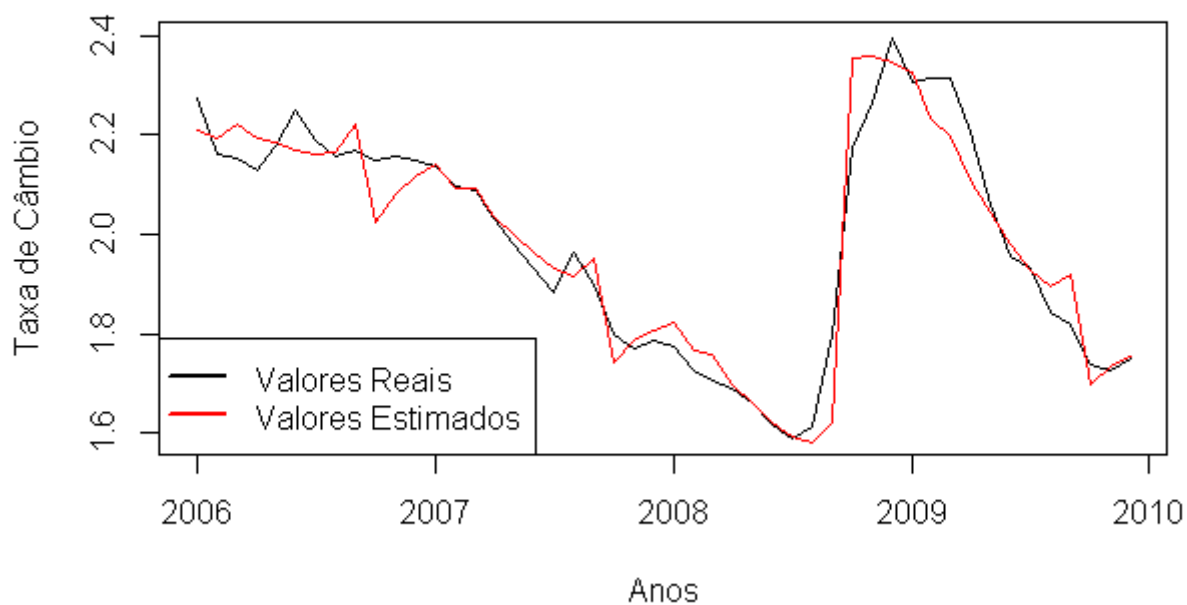
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 7,3031 - 2,4706t - 2,8063t^2 + 2,1611t^3$$

Já para o pós-crise teremos

$$\hat{y}_t = 15,0577 + 2,4706t - 2,8063t^2 + 2,1611t^3$$

Figura 11 – Modelo 2 da Taxa de Câmbio Ajustado



O ajuste foi bem semelhante do Modelo 1, e nosso modelo se reduziu apenas a uma regressão múltipla, pois apresenta apenas a componente da tendência.

3.2.4 Análise dos Resíduos e Adequabilidade dos Modelos da Taxa de Câmbio

As figuras 12 e 13 representam os resíduos dos modelos para a Taxa de Câmbio

Figura 12 – Resíduos do Modelo 1

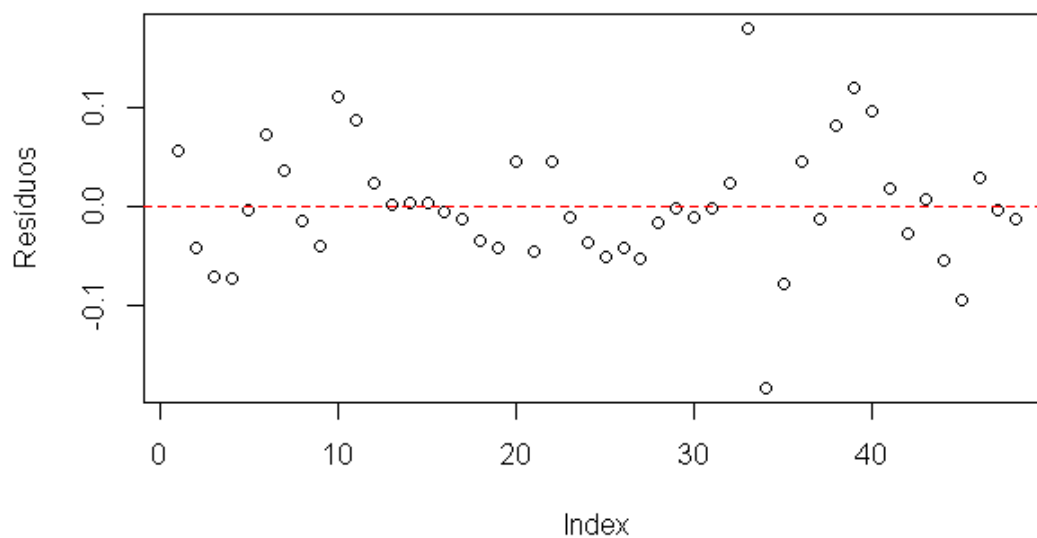
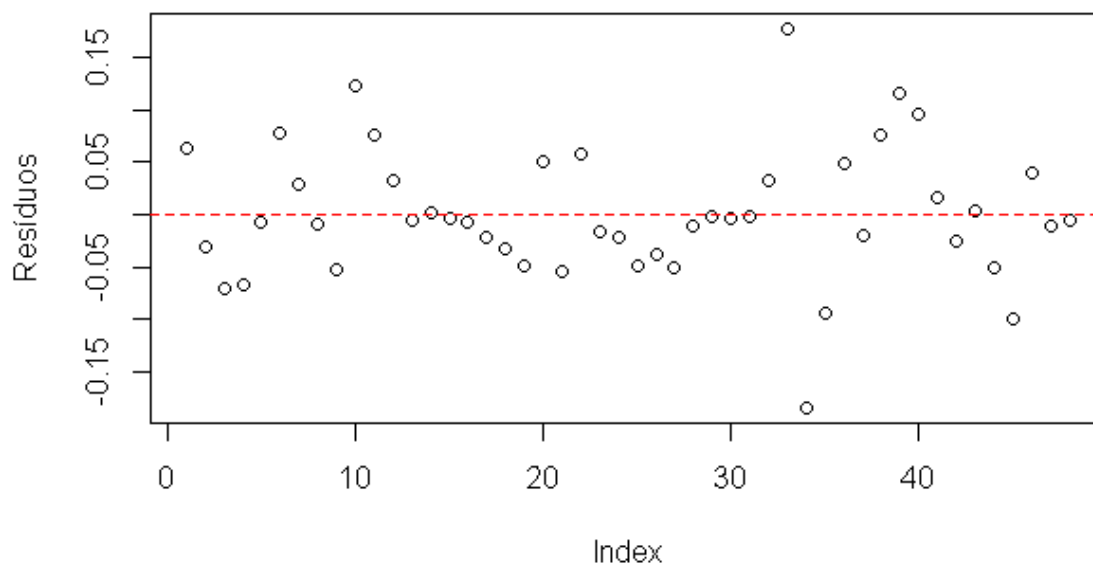


Figura 13 – Resíduos do Modelo 2



Como os dois modelos para a Taxa de Câmbio ficaram muito semelhantes então aparentemente os resíduos também possuem praticamente os mesmos gráficos, e analisando

não é possível afirmar que os resíduos possuem autocorrelação, faremos o teste de Durbin-Watson para isso

Tabela 13 – Teste de Durbin-Watson

	<i>Estatística</i>	<i>dl</i>	<i>du</i>
<i>Modelo 1</i>	$DW = 1,7245$	1,4064	1,671
<i>Modelo 2</i>	$DW = 1,8052$	1,362	1,721

Temos na tabela a estatística do teste e os limites inferiores e superiores tabelados para um $\alpha = 0,05$ para realizar o teste. Como em ambos os modelos $DW > du$ e $4 - DW > du$ então não rejeitamos a hipótese nula de que $\rho = 0$ e concluímos que não existe autocorrelação significativa nos resíduos.

É necessário também verificarmos a normalidade dos resíduos, a seguir temos a tabela com a estatística e o p-valor do teste de Shapiro-Wilk.

Tabela 14 – Teste de Shapiro-Wilk

	<i>Estatística</i>	<i>p – valor</i>
<i>Modelo 1</i>	$W = 0,96403$	0,1468
<i>Modelo 2</i>	$W = 0,96675$	0,1886

Pela Tabela 14 vemos que os resíduos apresentaram $p - valor > 0,05$, dessa forma não rejeitamos a hipótese nula a um nível de significância $\alpha = 0,05$ e concluímos que os resíduos seguem uma distribuição normal.

Com isso podemos assumir que nosso modelo está adequado, agora precisamos verificar a acurácia destes modelos.

Para isso vamos usar medidas como o RMSE, o MAE e o R^2 .

Tabela 15 – Medidas de Diagnóstico

	<i>RMSE</i>	<i>MAE</i>	R^2
<i>Modelo 1</i>	0,06129	0,21177	0,924
<i>Modelo 2</i>	0,6227	0,2141	0,922

Vemos que ambos os modelos obtiveram medidas muito semelhantes tanto para as medidas de erro quanto para o coeficiente de determinação que explicou cerca de 92% da variabilidade da Taxa de Câmbio. Ao analisar as medidas tem-se que o Modelo 1 obteve um melhor desempenho.

3.2.5 Modelagem do IPCA

Os parâmetros foram estimados pelo método de mínimos quadrados generalizados e foram utilizados apenas aqueles que foram significativos a um nível de significância de $\alpha = 0,05$.

As estimativas de mínimos quadrados generalizados desses parâmetros com $\hat{\rho} = -0,3$ para o Modelo 1 usando harmônicos da série do IPCA estão na Tabela 16.

Tabela 16 – Estimação dos Parâmetros do Modelo 1 do IPCA

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	113,7665	$2,397 \times 10^1$	$< 2 \times 10^{-16}$
β_1	-3,87	$6,114 \times 10^{-1}$	$5,45 \times 10^{-4}$
β_2	3,6741	$5,2776 \times 10^{-2}$	$9,27 \times 10^{-4}$
α_1	-5,403	1,0326	$7,465 \times 10^{-6}$
γ_1	4,8842	1,1215	$3,229 \times 10^{-5}$
θ	6,7688	$2,5782 \times 10^1$	$1,663 \times 10^{-7}$
δ	-6,7552	$7,5727 \times 10^{-1}$	$1,726 \times 10^{-7}$

Para o IPCA temos tanto uma tendência representada por β_0 , β_1 e β_2 quanto uma sazonalidade representada pelos parâmetros harmônicos α_1 e γ_1 , além disso ambas as variáveis “dummy” foram significativas para o Modelo 1.

O Modelo 1 do IPCA final é dado por:

$$\hat{y}_t = 113,7665 - 3,87t + 3,6741t^2 - 5,403 \cos\left(\frac{2\pi t}{12}\right) + 4,8842 \sin\left(\frac{8\pi t}{12}\right) + 6,7688Z_{it} - 6,7552Z_{it}t$$

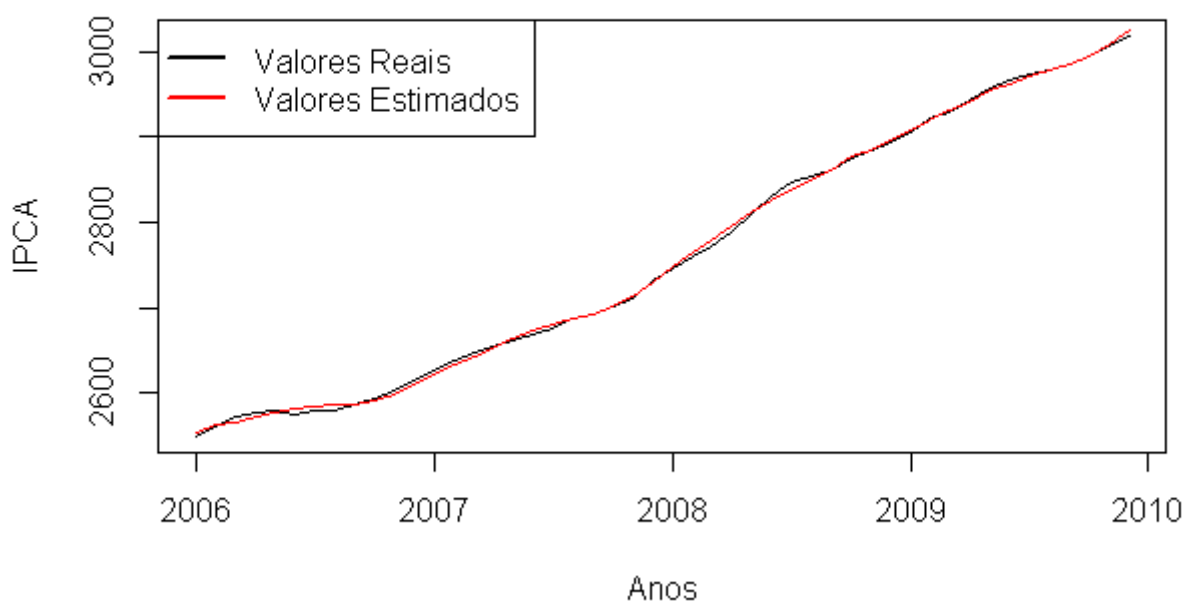
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 106,9977 + 2,8852t + 3,6741t^2 - 5,403 \cos\left(\frac{2\pi t}{12}\right) + 4,8842 \sin\left(\frac{8\pi t}{12}\right)$$

Já para o pós-crise teremos

$$\hat{y}_t = 120,5353 - 10,6252t + 3,6741t^2 - 5,403 \cos\left(\frac{2\pi t}{12}\right) + 4,8842 \sin\left(\frac{8\pi t}{12}\right)$$

Figura 14 – Modelo 1 do IPCA Ajustado



Percebe-se que o IPCA aparentemente não sofreu alteração com a crise, mantendo-se em constante crescimento, entretanto se olharmos para o Modelo 1 ele possui um bom ajuste e olhando para as variáveis “dummy” ele nos diz que houve diferença tanto na inclinação quanto no intercepto.

As estimativas dos parâmetros para o Modelo 2 com $\hat{\rho} = -0,95$ do IPCA estão na Tabela 17.

Tabela 17 – Estimação dos Parâmetros do Modelo 2 do IPCA

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	172,6772	$16,1974 \times 10^{-1}$	$< 2 \times 10^{-16}$
β_1	-2,6584	$6,51 \times 10^{-1}$	$1,21 \times 10^{-2}$
β_3	18,4717	$2,103 \times 10^{-4}$	$< 2 \times 10^{-16}$
θ	17,4992	14,609	$< 2 \times 10^{-16}$
δ	-16,8773	$4,443 \times 10^{-1}$	$< 2 \times 10^{-16}$

Diferente do Modelo 1 vemos que a série não possui a componente sazonal apenas a tendência, porém neste caso também temos as variáveis “dummy” significativas com $\delta < 0$, ou seja, houve diferença na inclinação após a crise, isso pode indicar que a taxa de crescimento da série diminuiu.

O Modelo 2 do IPCA final é dado por:

$$\hat{y}_t = 172,6772 - 2,6584t + 18,4717t^3 + 17,4992Z_{it} - 16,8773Z_{it}t$$

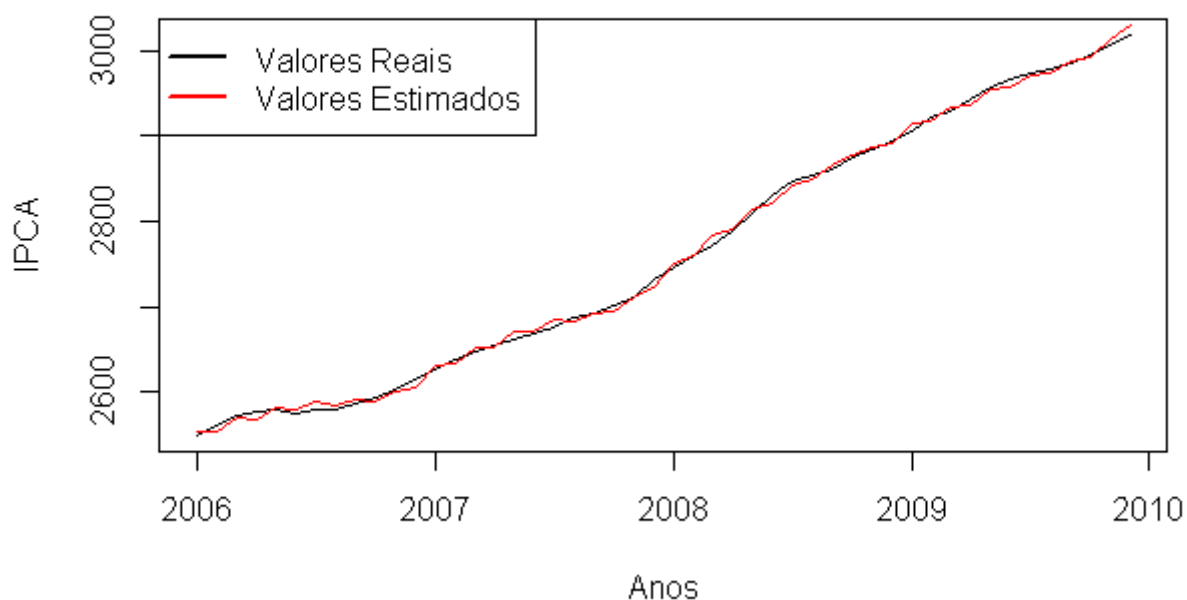
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 155,178 + 14,2189t + 18,4717t^3$$

Já para o pós-crise teremos

$$\hat{y}_t = 190,1764 - 19,5357t + 18,4717t^3$$

Figura 15 – Modelo 2 do IPCA Ajustado



O Modelo 2 obteve um bom ajuste, conseguindo se adequar bem a série ao longo do período estudado.

3.2.6 Análise dos Resíduos e Adequabilidade dos Modelos do IPCA

As figuras 16 e 17 representam os resíduos dos modelos para o IPCA

Figura 16 – Resíduos do Modelo 1

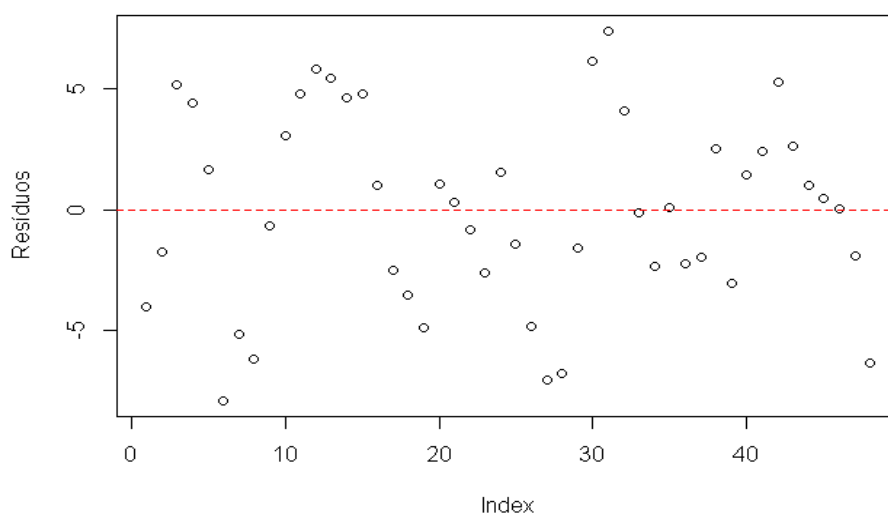
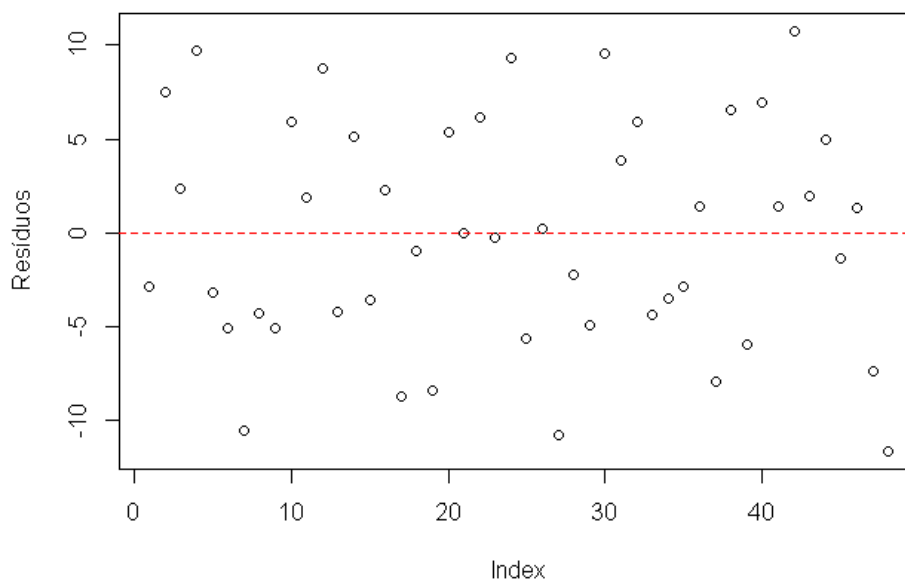


Figura 17 – Resíduos do Modelo 2



Graficamente os resíduos não aparentam ter uma autocorrelação muito forte, o gráfico dos resíduos do Modelo 1 apresenta uma certa sazonalidade já o Modelo 2 não, para confirmar isso faremos o teste de Durbin-Watson.

Tabela 18 – Teste de Durbin-Watson

	<i>Estatística</i>	<i>dl</i>	<i>du</i>
<i>Modelo 1</i>	$DW = 0,7693$	1,271	1,826
<i>Modelo 2</i>	$DW = 1,9736$	1,362	1,721

Temos na tabela a estatística do teste e os limites inferiores e superiores tabelados para um $\alpha = 0,05$ para realizar o teste. Como no Modelo 1 $DW < dl$ então rejeitamos a hipótese nula de que $\rho = 0$ e concluímos que existe autocorrelação significativa nos resíduos, neste caso ela é positiva. Para o Modelo 2 temos que $DW > du$ e $4 - DW < du$, dessa forma não rejeitamos a hipótese nula e concluímos que não existe autocorrelação.

É necessário também verificarmos a normalidade dos resíduos, a seguir temos a tabela com a estatística e o p-valor do teste de Shapiro-Wilk.

Tabela 19 – Teste de Shapiro-Wilk

	<i>Estatística</i>	<i>p – valor</i>
<i>Modelo 1</i>	$W = 0,97511$	0,3948
<i>Modelo 2</i>	$W = 0,97042$	0,263

Pela Tabela 19 vemos que os resíduos apresentaram $p - valor > 0,05$, dessa forma não rejeitamos a hipótese nula a um nível de significância $\alpha = 0,05$ e concluímos que os resíduos seguem uma distribuição normal.

Com isso podemos assumir que nosso modelo está adequado, agora precisamos verificar a acurácia destes modelos.

Para isso vamos usar medidas como o RMSE, o MAE e o R^2 .

Tabela 20 – Medidas de Diagnóstico

	<i>RMSE</i>	<i>MAE</i>	R^2
<i>Modelo 1</i>	3,9317	1,8082	0,9984
<i>Modelo 2</i>	6,0015	2,2616	0,9984

Ambos os modelos explicam a mesma quantidade da variabilidade do IPCA. O Modelo 1 possui medidas de erro menor se comparado ao Modelo 2, porém como o Modelo 1 não passou no teste de autocorrelação dos resíduos o Modelo 2 é mais adequado.

3.2.7 Modelagem do Índice de Preços das Importações

As dos parâmetros para o Modelo 1 usando harmônicos e $\hat{\rho} = -0,333$ da série do Índice de Preços das Importações estão na Tabela 21.

Tabela 21 – Estimação dos Parâmetros do Modelo 1 do Índice de Preços das Importações

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	37,101	4,923	$< 2 \times 10^{-16}$
β_1	-16,7671	$1,9712 \times 10^{-1}$	$< 2 \times 10^{-16}$
β_3	18,093	$6,3181 \times 10^{-5}$	$6,866 \times 10^{-3}$
α_1	-2,8963	$3,3448 \times 10^{-1}$	$7,465 \times 10^{-6}$
θ	22,4237	4,4537	$< 2 \times 10^{-16}$
δ	-23,6052	$1,35 \times 10^{-1}$	$< 2 \times 10^{-16}$

Para o Modelo 1 temos a tendência representada por β_0 , β_1 e β_3 e a sazonalidade representada pelo harmônico α_1 , além disso ambas as variáveis “dummy” foram significativas

para o Modelo 1, sendo $\delta < 0$ logo a inclinação da série diminuiu após a crise além do intercepto ter aumentado pois $\theta > 0$.

O Modelo 1 do Índice de Preços das Importações final é dado por:

$$\hat{y}_t = 37,101 - 16,7671t + 18,093t^3 - 2,8963 \cos\left(\frac{2\pi t}{12}\right) + 22,4237Z_{it} - 23,6052Z_{it}t$$

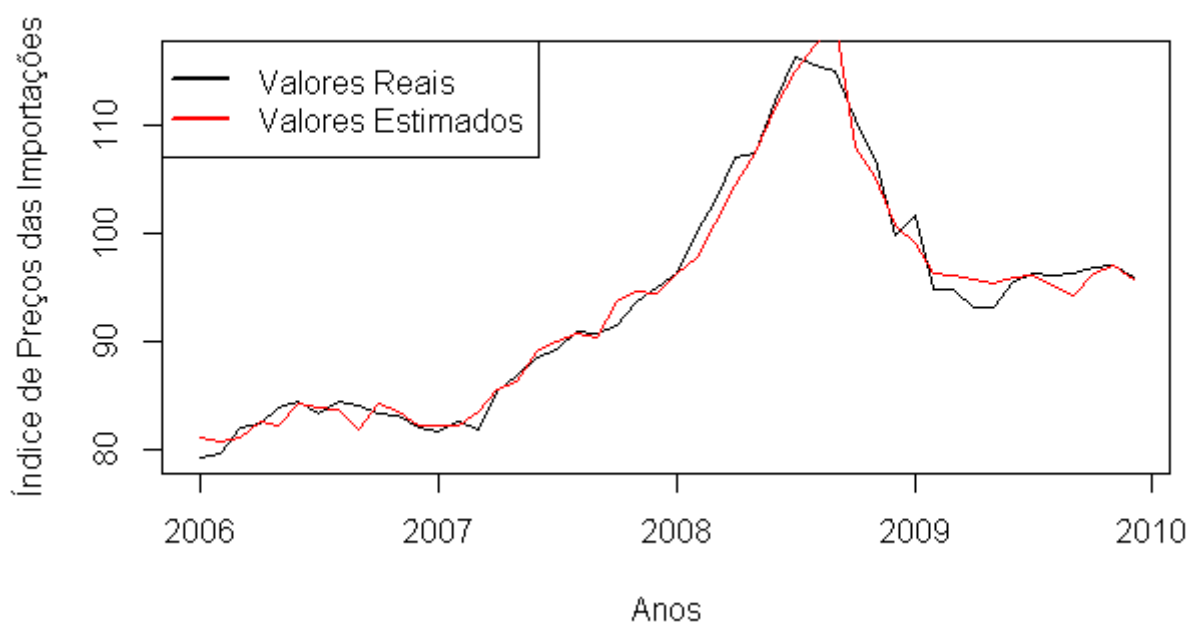
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 14,6773 + 6,8381t + 18,093t^3 - 2,8963 \cos\left(\frac{2\pi t}{12}\right)$$

Já para o pós-crise teremos

$$\hat{y}_t = 59,5247 - 40,3723t + 18,093t^3 - 2,8963 \cos\left(\frac{2\pi t}{12}\right)$$

Figura 18 – Modelo 1 do Índice de Preços das Importações Ajustado



A série vinha em uma crescente desde 2006 e com a crise em 09/2008 há uma queda brusca e o Modelo 1 conseguiu prever bem este comportamento, nos indicando que houve diferenças significativas do pré e pós crise

As estimativas dos parâmetros para o Modelo 2 com $\hat{\rho} = -0,95$ do estão na Tabela 22.

Tabela 22 – Estimação dos Parâmetros do Modelo 2 do Índice de Preços das Importações

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	40,8334	4,4864	$< 2 \times 10^{-16}$
β_1	-18,4911	$1,8032 \times 10^{-1}$	$< 2 \times 10^{-16}$
β_3	19,8393	$5,8254 \times 10^{-5}$	$< 2 \times 10^{-16}$
θ	24,7834	4,0465	$< 2 \times 10^{-16}$
δ	-26,0869	$1,228 \times 10^{-1}$	$< 2 \times 10^{-16}$

Vemos que a série não possui a componente sazonal apenas a tendência, e mais uma vez temos as variáveis “dummy” significativas com $\delta < 0$, ou seja, houve diferença na inclinação após a crise, isso é confirmado ao olharmos para a série e vemos uma queda a partir de 09/2008

O Modelo 2 do Índice de Preços das Importações final é dado por:

$$\hat{y}_t = 40,8334 - 18,4911t + 19,8393t^3 + 24,7834Z_{it} - 26,0869Z_{it}t$$

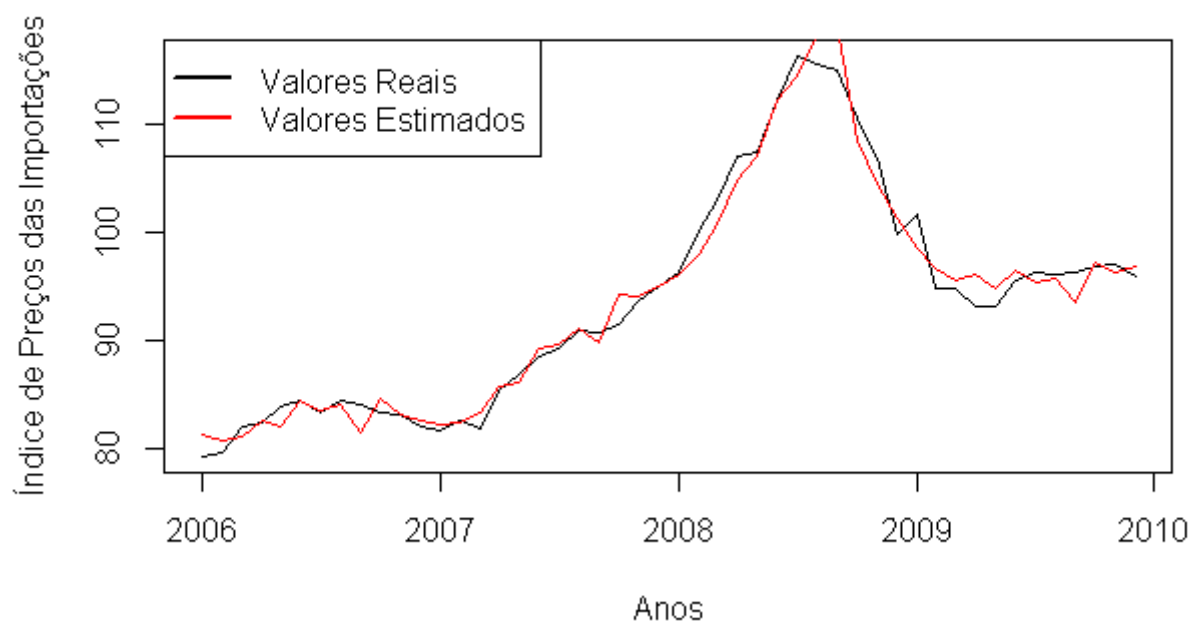
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 16,05 + 7,5958t + 19,8393t^3$$

Já para o pós-crise teremos

$$\hat{y}_t = 65,6168 - 44,578t + 19,8393t^3$$

Figura 19 – Modelo 2 do Índice de Preços das Importações Ajustado



3.2.6 Análise dos Resíduos e Adequabilidade dos Modelos do Índice de Preços das Importações

As figuras 20 e 21 representam os resíduos dos modelos para o Índice de Importações.

Figura 20 – Resíduos do Modelo 1

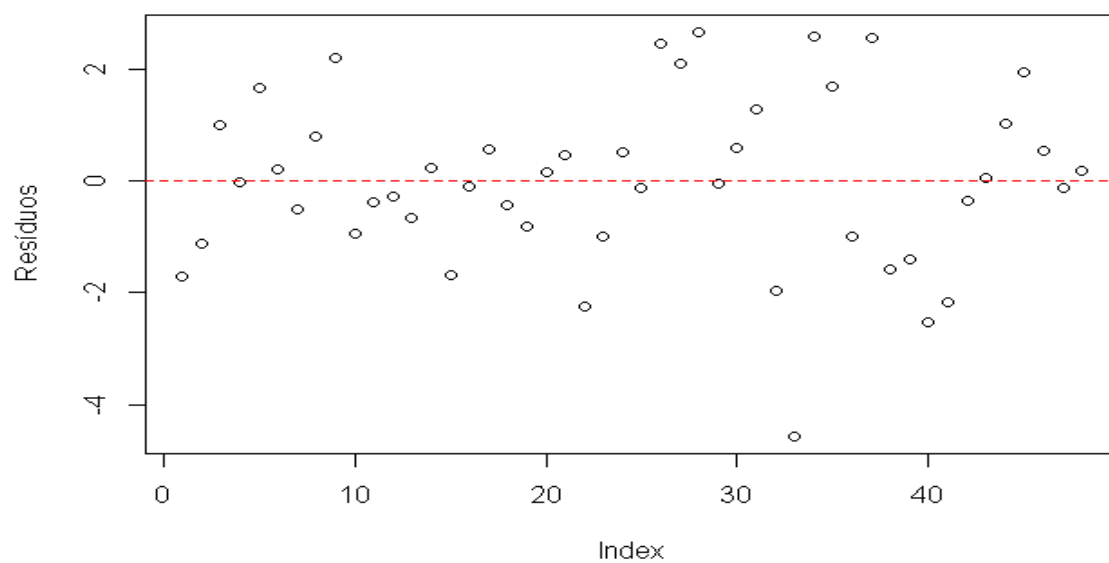
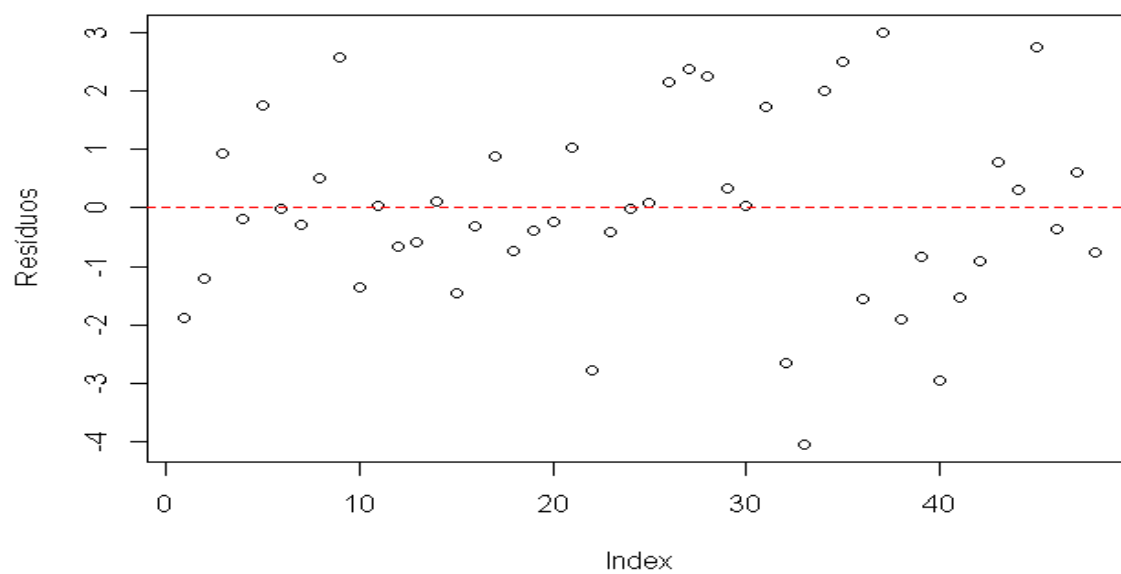


Figura 21– Resíduos do Modelo 2



Graficamente os resíduos não aparentam ter uma autocorrelação muito forte, o gráfico dos resíduos do Modelo 1 possui um ponto discrepante, já o Modelo 2 apresenta um comportamento que não é possível afirmar se os resíduos são autocorrelacionados, para confirmar isso faremos o teste de Durbin-Watson.

Tabela 23 – Teste de Durbin-Watson

	<i>Estatística</i>	<i>dl</i>	<i>du</i>
<i>Modelo 1</i>	$DW = 1,6412$	1,317	1,772
<i>Modelo 2</i>	$DW = 1,8793$	1,362	1,721

Temos na tabela a estatística do teste e os limites inferiores e superiores tabelados para um $\alpha = 0,05$ para realizar o teste. Como no Modelo 1 $dl < DW < du$ então o teste é inconclusivo e não é capaz de rejeitar ou não hipótese nula de que $\rho = 0$. Para o Modelo 2 temos que $DW > du$ e $4 - DW < du$, dessa forma não rejeitamos a hipótese nula e concluímos que não existe autocorrelação.

É necessário também verificarmos a normalidade dos resíduos, a seguir temos a tabela com a estatística e o p-valor do teste de Shapiro-Wilk.

Tabela 24 – Teste de Shapiro-Wilk

	<i>Estatística</i>	<i>p – valor</i>
<i>Modelo 1</i>	$W = 0,97314$	0,3341
<i>Modelo 2</i>	$W = 0,97725$	0,47

Pela Tabela 24 vemos que os resíduos apresentaram $p - valor > 0,05$, dessa forma não rejeitamos a hipótese nula a um nível de significância $\alpha = 0,05$ e concluímos que os resíduos seguem uma distribuição normal.

Com isso podemos assumir que nosso modelo está adequado, agora precisamos verificar a acurácia destes modelos.

Para isso vamos usar medidas como o RMSE, o MAE e o R^2 .

Tabela 25 – Medidas de Diagnóstico

	<i>RMSE</i>	<i>MAE</i>	R^2
<i>Modelo 1</i>	1,4971	1,0712	0,9783
<i>Modelo 2</i>	1,5831	1,1061	0,9757

O coeficiente de determinação de ambos está bem próximo com cerca de 97% da variabilidade da variável repostada sendo explicada. Vemos que o Modelo 1 possui medidas de erro menor se comparado ao Modelo 2, porém como o Modelo 1 não conseguiu ter um teste conclusivo para a autocorrelação dos resíduos, o Modelo 2 aparenta ser mais adequado.

3.2.7 Modelagem do Índice de Preços das Exportações

Os parâmetros foram estimados pelo método de mínimos quadrados generalizados e foram utilizados apenas aqueles que foram significativos a um nível de significância de $\alpha = 0,05$.

As dos parâmetros para o Modelo 1 usando harmônicos e $\hat{p} = -0,01$ da série do Índice de Preços das Exportações estão na Tabela 26.

Tabela 26 – Estimação dos Parâmetros do Modelo 1 do Índice de Preços das Exportações

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	20,7389	$1,1102 \times 10^1$	$< 2 \times 10^{-16}$
β_1	-11,8467	$2,8542 \times 10^{-1}$	$7,687 \times 10^{-13}$
β_2	-4,841	$2,3698 \times 10^{-2}$	$3,6488 \times 10^{-5}$
β_3	7,8295	$4,3630 \times 10^{-4}$	$9,7609 \times 10^{-9}$

γ_1	-3,5173	$5,1628 \times 10^{-1}$	$1,4104 \times 10^{-3}$
θ	13,5526	$1,1856 \times 10^1$	$2,5313 \times 10^{-14}$
δ	-14,4818	$3,4827 \times 10^{-1}$	$4,4409 \times 10^{-15}$

No Modelo 1 temos quatro parâmetros significativos para tendência, β_0 , β_1 , β_2 e β_3 e a sazonalidade possui apenas γ_1 , as variáveis “dummy” também foram significativas, sendo $\delta < 0$ logo a inclinação da série diminuiu após a crise além do intercepto ter aumentado pois $\theta > 0$.

O Modelo 1 do Índice de Preços das Exportações final é dado por:

$$\hat{y}_t = 20,7389 - 11,8467t - 4,841t^2 + 7,8295t^3 - 3,5173 \sin\left(\frac{2\pi t}{12}\right) + 13,5526Z_{it} - 14,4818Z_{it}t$$

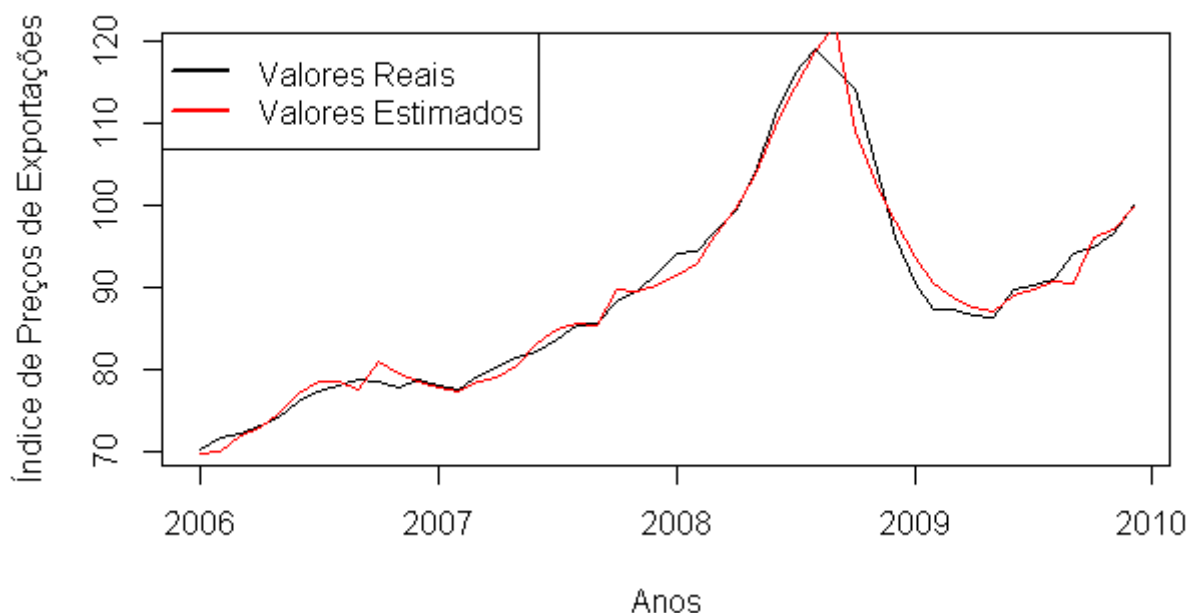
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 7,1863 + 2,6351t - 4,841t^2 + 7,8295t^3 - 3,5173 \sin\left(\frac{2\pi t}{12}\right)$$

Já para o pós-crise teremos

$$\hat{y}_t = 34,2915 - 26,3285t - 4,841t^2 + 7,8295t^3 - 3,5173 \sin\left(\frac{2\pi t}{12}\right)$$

Figura 22 – Modelo 1 do Índice de Preços das Exportações Ajustado



A série vinha em uma crescente desde 2006 e com a crise em 09/2008 há uma queda brusca pela desaceleração da economia dos países Europeus e os EUA dessa forma os tanto o volume quanto os preços de exportação diminuíram principalmente os manufaturados.

O Modelo 1 conseguiu se ajustar bem a este comportamento, indicando diferenças significativas do pré e pós crise.

As estimativas dos parâmetros para o Modelo 2 com $\hat{\rho} = -0,90$ do estão na Tabela 27.

Tabela 27 – Estimação dos Parâmetros do Modelo 2 do Índice de Preços das Exportações

<i>Parâmetro</i>	<i>Estimativa</i>	<i>Desvio Padrão</i>	<i>P – valor</i>
β_0	26,8584	7,068	$< 2 \times 10^{-16}$
β_1	-13,5249	$2,8402 \times 10^{-1}$	$8,8817 \times 10^{-15}$

β_3	15,1137	$9,1704 \times 10^{-5}$	$4,409 \times 10^{-16}$
θ	17,7743	6,3769	$< 2 \times 10^{-16}$
δ	-19,0673	$1,9350 \times 10^{-1}$	$< 2 \times 10^{-16}$

A série não possui a componente sazonal apenas a tendência, temos as variáveis “dummy” significativas com $\delta < 0$, o que indica diferença na inclinação após a crise.

O Modelo 2 do Índice de Preços das Exportações final é dado por:

$$\hat{y}_t = 26,8584 - 13,5249t + 15,1137t^3 + 17,7743Z_{it} - 19,0673Z_{it}t$$

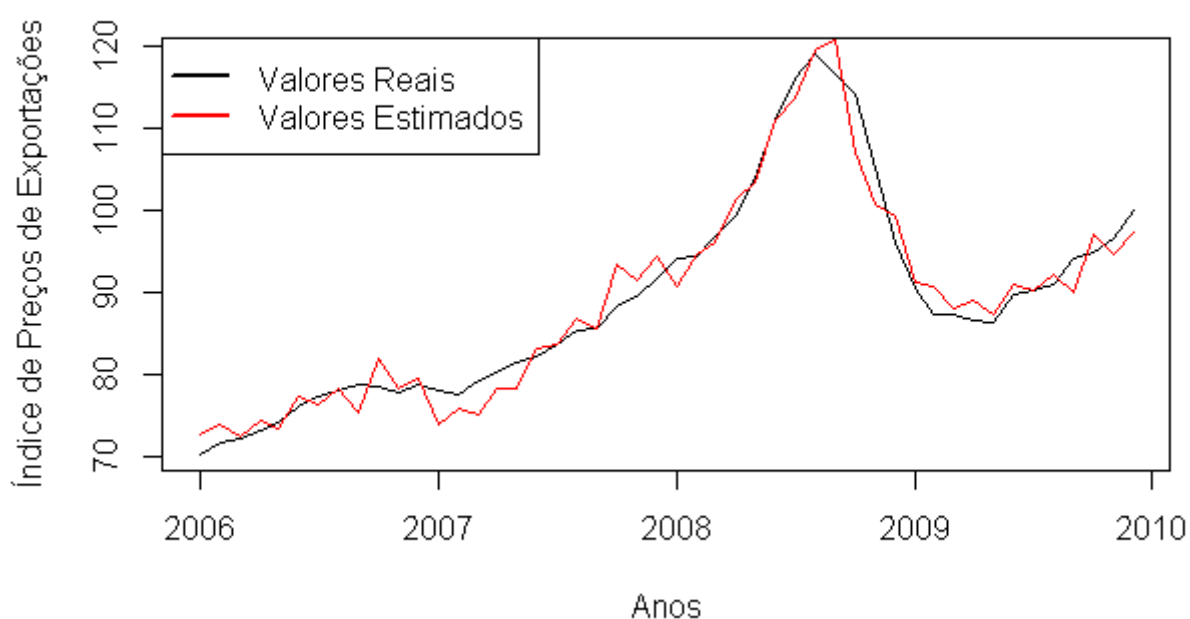
Para o período pré-crise teremos o seguinte modelo

$$\hat{y}_t = 9,0841 + 5,5424t + 15,1137t^3$$

Já para o pós-crise teremos

$$\hat{y}_t = 44,6327 + 32,5922t + 15,1137t^3$$

Figura 23 – Modelo 2 do Índice de Preços das Exportações Ajustado



O ajuste do Modelo 2 está bom também, porém aparentemente o Modelo 1 se ajustou melhor, olhando para o início da série o Modelo 2 teve dificuldade de se adequar, entretanto em 09/2008 conseguiu se ajustar melhor ao pico que o Modelo 1

3.2.6 Análise dos Resíduos e Adequabilidade dos Modelos do Índice de Preços das Exportações.

As figuras 24 e 25 representam os resíduos dos modelos para o Índice dos Preços de Exportações.

Figura 24– Resíduos do Modelo 1

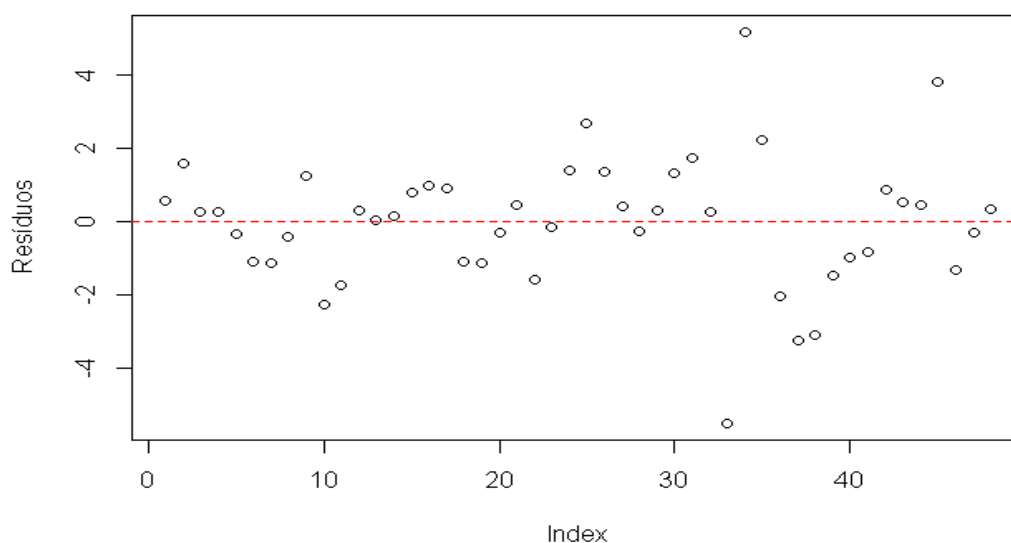
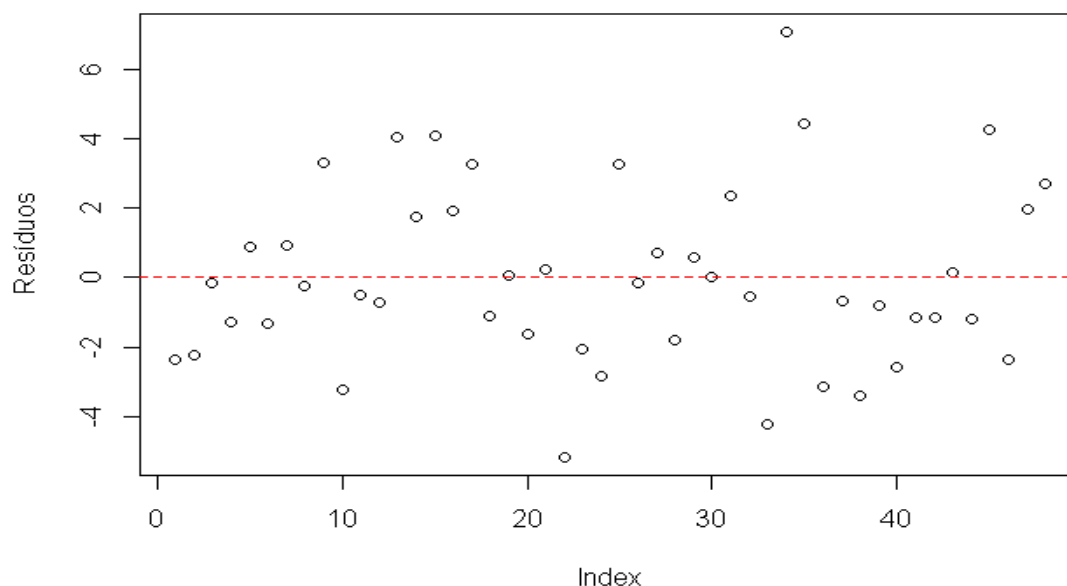


Figura 25 – Resíduos do Modelo 2



Graficamente os resíduos não aparentam ter uma autocorrelação muito forte, o gráfico do Modelo 1 possui alguns pontos mais dispersos, porém aparentam estar em torno do 0, o Modelo 2 também parece apresentar resíduos não autocorrelacionados para confirmar isso faremos o teste de Durbin-Watson.

Tabela 28 – Teste de Durbin-Watson

	<i>Estatística</i>	<i>dl</i>	<i>du</i>
<i>Modelo 1</i>	$DW = 1,8183$	1,27	1,826
<i>Modelo 2</i>	$DW = 1,9333$	1,362	1,721

Temos na tabela a estatística do teste e os limites inferiores e superiores tabelados para um $\alpha = 0,05$ para realizar o teste. Como no Modelo 1 $dl < DW < du$ por muito pouco então o teste é inconclusivo e não é capaz de rejeitar ou não hipótese nula de que $\rho = 0$. Para o Modelo 2 temos que $DW > du$ e $4 - DW < du$, dessa forma não rejeitamos a hipótese nula e concluímos que não existe autocorrelação.

É necessário também verificarmos a normalidade dos resíduos, a seguir temos a tabela com a estatística e o p-valor do teste de Shapiro-Wilk.

Tabela 29 – Teste de Shapiro-Wilk

	<i>Estatística</i>	<i>p – valor</i>
<i>Modelo 1</i>	$W = 0,959$	0,0919
<i>Modelo 2</i>	$W = 0,974$	0,3603

Pela Tabela 29 vemos que os resíduos apresentaram $p - valor > 0,05$, dessa forma não rejeitamos a hipótese nula a um nível de significância $\alpha = 0,05$ e concluímos que os resíduos seguem uma distribuição normal.

Com isso podemos assumir que nosso modelo está adequado, agora precisamos verificar a acurácia destes modelos.

Para isso vamos usar medidas como o RMSE, o MAE e o R^2 .

Tabela 30 – Medidas de Diagnóstico

	<i>RMSE</i>	<i>MAE</i>	R^2
<i>Modelo 1</i>	1,7483	1,1235	0,98
<i>Modelo 2</i>	2,5341	1,4151	0,9541

Como o Modelo 1 não conseguiu ter um teste conclusivo para a autocorrelação dos resíduos, o Modelo 2 mesmo apresentando medidas de erro maior e um coeficiente de determinação menor, ainda sim é mais adequado.

4 CONCLUSÕES

Baseando-se nas análises feitas, é possível concluir que a crise de 2008 promoveu uma mudança em todos os indicadores estudados durante o período estudado (2006-2009). Vemos que durante o período de 2006 até 2008 indicadores como a Taxa de Juros e de Câmbio vinham caindo por conta do crescimento da inflação, já os indicadores como IPCA e Índices de Preços de Importações e Exportações vinham crescendo, com o estopim da crise em setembro de 2008 ocorre um aumento nessas Taxas de Juros e Câmbio pela desvalorização do dólar em relação ao real, entretanto com a ajuda do governo injetando dinheiro na economia houve um aumento na liquidez, resultando assim na diminuição da Taxa de Juros de forma gradual. Com a diminuição da Taxa de Juros a inflação aumenta e com isso o real desvalorizou em relação ao dólar, fazendo assim a Taxa de Câmbio diminuir também.

Com relação aos índices a crise teve impactos diferentes, no IPCA esta alteração não foi grande, porém com os Preços de Importações e Exportações houve mudança, pois com a crise os países desenvolvidos da Europa e EUA tiveram uma desaceleração em suas economias fazendo os preços diminuírem de forma geral, no entanto pouco tempo após a crise esses preços voltaram a aumentar.

Através dos modelos obtidos foi possível acompanhar essas variações, vemos que houve diferença significativa tanto na inclinação da reta quanto no deslocamento em todos os casos, o que significa que a crise teve um certo impacto mesmo que de forma pequena, os modelos obtiveram excelentes ajustes e com R^2 bem elevados.

Para o IPCA houve problemas com autocorrelação nos resíduos mesmo utilizando a estimação de mínimos quadrados generalizados é possível notar uma certa tendência e o Modelo 1(usando harmônicos) não passou no teste de Durbin-Watson, já no Índice de Preços das Importações e Exportações o teste de Durbin-Watson foi inconclusivo para o Modelo 1 (usando harmônicos) pois a estatística do teste está entre os limites inferiores e superiores, ou seja, em uma zona de inconclusão do teste, assim para estes indicadores os modelos utilizando a matriz de delineamento se saiu melhor.

Ambos os modelos se saíram bem para as Taxas de Juros e de Câmbio, passando em todos os testes e obtendo boas métricas, porém ao olharmos para as métricas RMSE e MAE temos que os modelos harmônicos se saíram melhor por uma pequena margem.

No geral os modelos utilizando a matriz de delineamento obtiveram melhores resultados com relação a autocorrelação dos resíduos e não tiveram ajustes tão inferiores ao dos modelos

usando harmônico, entretanto os modelos foram simplificados devido à falta de significância estatística na maioria das variáveis em relação à componente sazonal. A única exceção foi a Taxa de Juros, que mostrou um parâmetro significativo para a componente sazonal. Assim, os modelos foram reduzidos a regressões lineares múltiplas com variáveis 'dummy', destacando a importância apenas da componente de tendência, que foi a única significativa.

REFERÊNCIAS

BUENO, R. D'LOSSO S. **Econometria de Séries Temporais**. CENGAGE Learning, São Paulo, 2008.

DRAPER, N.R.; SMITH, H. **Applied Regression Analysis**. 4a. ed. New York, John Wiley & Sons, 1968

.

FGV. Disponível em: <https://portalibre.fgv.br/>

FILGUEIRAS, Luiz. **A crise do mercado imobiliário dos EUA e suas implicações para a economia brasileira**. Revista Conjuntura & Planejamento Econômico, Salvador, n. 156, p. 88-93, jul./set. 2007

GUJARATI, D. N. **Econometria Básica**. São Paulo, Makron Books LTDA, 2000

HOFFMAN, R. e VIERA, S. **Análise de Regressão – Uma Introdução à Econometria**. 2a ed., Editora Hucitec Ltda, São Paulo, 1987.

IBGE. Disponível em: <https://www.ibge.gov.br/>

IPEA. Disponível em: <http://www.ipeadata.gov.br/>

MACKINNON, J.G. Numerical Distribution Functions for Unit Root and Cointegration Tests. **Journal of Applied Econometrics**, 11, 1996.

Mazuqueli, J.R. **Impacto da Covid nos Principais Indicadores Econômicos do Brasil**. Faculdade de Estatística, Universidade Estadual Paulista (Unesp). 2022. Disponível em: <http://hdl.handle.net/11449/217367>

MONTGOMERY, D.C.; PECK, E.A.; VINING, G.G. **Introduction to Linear Regression Analysis**. 4a. ed, Hoboken, John Wiley & Sons, 2006.

MORETTIN, P.A. & TOLOI, C.M. **Análise de Séries Temporais**. São Paulo, Edgard Blücher, 2004.

Pires, S.S. **A crise financeira internacional de 2008 e seus desdobramentos sobre a economia brasileira**. VI Jornada Internacional de Políticas Públicas, 2013. Disponível em: <http://www.joinpp.ufma.br/jornadas/joinpp2013/JornadaEixo2013/anais-eixo1-mundializacaoestadosnacionaispoliticaspublicas/acrisefinanceirainternacionalde2008eseusdesdobramentossobreaeconomiabrasileira.pdf>

Santos, A.R. **Crise financeira mundial de 2008 e seus impactos na economia brasileira**. UNIVERSIDADE FEDERAL DO CEARÁ FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, ATUÁRIA E CONTABILIDADE CURSO DE CIÊNCIAS ECONÔMICAS, 2012. Disponível em: https://repositorio.ufc.br/bitstream/riufc/32525/1/2012_tcc_arsantos.pdf

APÊNDICE

A seguir é apresentado o código em R utilizado para modelar a Taxa de Juros utilizando harmônicos

```
##### JUROS #####
require(dplyr)
library(lmtest)
```

```

library(car)
library(lmtest)
library(caret)
library(tseries)
library(forecast)
library(DescTools)

JUROS <- read.csv("C:/ESRI/Roger/TCC/Indicadores/Juros.csv",sep = ';')
JUROS <- JUROS %>% filter(JUROS$Data > 2005.12 & JUROS$Data <=2009.12)

colnames(JUROS) <- c('Data','JUROS')

y <- JUROS$JUROS

adf.test(y)

TCTS= ts(y, frequency = 12, start = c(2006,1),end = c(2009,12))
plot(TCTS,xlab='Anos',ylab='Taxa de Juros')
abline(v = 2008.60, col = 'blue', lwd = 2)
legend( x = 'topright', legend = '09/2008', col = 'blue',
       lwd = 2, merge = TRUE)

pi = 3.1415
aux <- rep(1,48)
aux2 <- 1:48

T1 <- matrix(c(aux,aux2,aux2^2,aux2^3),ncol = 4,nrow=48)
T1

# Tamanho da matriz
num_linhas <- 48
num_colunas <- 12

# Inicialize a matriz com zeros
H <- matrix(0, nrow = num_linhas, ncol = num_colunas)

# Preencha a matriz de acordo com as especificações
for (i in 1:num_linhas) {
  for (k in 1:6) {
    H[i, (k)] <- cos(i * (2 * pi * k / 12))
    H[i, (k+6)] <- sin(i * (2 * pi * k / 12))
  }
}

# Exiba a matriz H

```

```
print(H)
```

```
dummy_m <- matrix(c(rep(-1,33),rep(1,15)),nrow = 48)
dummy_m
```

```
dummyt_m <- matrix(0, nrow = 48, ncol = 1)
for (i in 1:48) {
  dummyt_m[i,] <- i*dummy_m[i,]
}
```

```
dummyt_m
```

```
X <- cbind(T1,H,dummy_m,dummyt_m)
X
```

```
rho <- 0.75 # único rho que deu certo
```

```
Y <- matrix(y,ncol = 1)
Y
```

```
# Tamanho da matriz (número de linhas)
n_matriz <- 48
```

```
# Inicializar a matriz com zeros
matriz <- matrix(0, nrow = n_matriz, ncol = n_matriz)
```

```
# Preencher a diagonal principal com 1
diag(matriz) <- 1
```

```
# Preencher a posição à esquerda da diagonal com -rho
for (i in 2:n_matriz) {
  matriz[i, i - 1] <- -rho
}
```

```
matriz[1,1] <- sqrt(1-rho^2)
```

```
lambda <- matriz
```

```
Vinv <- t(lambda) %*% lambda
Vinv
```

```
betas <- solve(t(X) %*% Vinv %*% X) %*% t(X) %*% Vinv %*% Y
betas
```

```
y_est <- X %*% betas
y_est
```

```
erros <- y-y_est
erros
sum(erros)
```

```
plot(erros,ylab='Resíduos');abline(h=0,col='red',lty=2)
```

```
p <- 18
n <- 48
```

```
#ESTIMATIVA DE SIGMA^2
s2 <- (t(y)%*%Vinv%*%Y) - (t(betas)%*%t(X)%*%Vinv%*%Y)
s2 <- s2/(n-p)
s2
```

```
C <- solve(t(X) %*% Vinv %*% X) * 0.007126499
C
```

```
# VENDO O AJUSTE DO MODELOS
```

```
yajustado <- y_est
a <- ts(yajustado, frequency = 12, start = c(2006, 01),end = c(2009, 12))
b <- ts(y, frequency = 12, start = c(2006, 01),end = c(2009, 12))
plot(b,ylab='Taxa de Juros',xlab='Anos')
lines(a,col='red')
legend( x = 'topright', legend = c('Valores Reais ','Valores Estimados'), col = c('red','blue'),
       lwd = 2, merge = TRUE)
```

```
# Inicialize vetores para armazenar as estatísticas t e p-valores
t_statistics <- numeric(p)
p_values <- numeric(p)
```

```
# Loop para calcular estatísticas t e p-valores para cada parâmetro
for (i in 1:p) {
  coef_i[i] <- betas[i] # Coeficiente do parâmetro i
  erro_padrao_i[i] <- sqrt(C[i,i]) # Erro padrão do coeficiente i
  graus_de_liberdade_i <- n - p # Graus de liberdade
```

```
# Calcula a estatística t e p-valor
t_statistics[i] <- coef_i[i] / erro_padrao_i[i]
```



```

p_values[i] <- 2 * (1 - pt(abs(t_statistics[i]), df = graus_de_liberdade_i))
}

nomes <- c("Beta_0","Beta_1","Beta_2","Beta_3","Cosseno_1","Cosseno_2",
          "Cosseno_3","Cosseno_4","Cosseno_5","Cosseno_6","Seno_1","Seno_2",
          "Seno_3","Seno_4","Seno_5","Seno_6","Dummy","Dummyt")
# Exibe os resultados
resultados <- data.frame(Parâmetro = nomes, T_Statistic = t_statistics,ep = erro_padrao_i,
P_Value = p_values)
print(resultados)

qqnorm(erros)
qqline(erros,col='red')

plot(erros,ylab='Resíduos');abline(h=0,col='red',lty=2)

shapiro.test(erros)# TESTE DE NORMALIDADE DOS RESÍDUOS

durbinWatsonTest(model=erros[,1])

4-durbinWatsonTest(model=erros[,1])

# Diagnóstico do Modelo
RMSE = function(m, o) sqrt(mean((m - o)^2))
MAE = function(m, o) sqrt(mean(abs(m - o)))
RMSE(yajustado,y)
MAE(yajustado,y)

# Coeficiente de determinação
observacoes_reais <- y
observacoes_previstas <- yajustado

# Cálculo da média das observações reais
media_observacoes_reais <- mean(observacoes_reais)

# Cálculo da soma total dos quadrados (SST)
sst <- sum((observacoes_reais - media_observacoes_reais)^2)

# Cálculo do erro residual (SSE)
sse <- sum((observacoes_reais - observacoes_previstas)^2)

# Cálculo do coeficiente de determinação (R-squared)
r_squared <- 1 - (sse / sst)

```

```
# Imprime o valor do R-squared  
cat("O coeficiente de determinação (R-squared) é:", r_squared, "\n")
```