

**FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO**

ARTHUR PEREIRA GOES

MODELOS ROBUSTOS DE PREVISÃO DE INFLAÇÃO

São Paulo
2019

ARTHUR PEREIRA GOES

MODELOS ROBUSTOS DE PREVISÃO DE INFLAÇÃO

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getúlio Vargas como requisito para obtenção do título de mestre em Economia

Orientador: Emerson Fernandes Marçal

São Paulo
2019

Goes, Arthur Pereira.

Modelos robustos de previsão de inflação / Arthur Pereira Goes. - 2019.
33 f.

Orientador: Emerson Fernandes Marçal.

Dissertação (mestrado profissional MPFE) – Fundação Getulio Vargas,
Escola de Economia de São Paulo.

1. Inflação - Brasil. 2. Inflação - Previsão. 3. Índice nacional de preços ao
consumidor amplo. 4. Modelos econométricos. I. Marçal, Emerson Fernandes.
II. Dissertação (mestrado profissional MPFE) – Escola de Economia de São
Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 336.748.12(81)

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-
010191/O

Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

ARTHUR PEREIRA GOES

MODELOS ROBUSTOS DE PREVISÃO DE INFLAÇÃO

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para obtenção do título de mestre em Economia

Data de aprovação: __/__/____

Banca examinadora:

Prof. Dr. Emerson Fernandes Marçal
FGV-EESP (Orientador)

Prof. Marcelo Kfoury
FGV-EESP

Prof. Diogo de Prince
UNIFESP

RESUMO

Esse estudo tem como objetivo avaliar diferentes modelos para a previsão do índice de inflação nacional (IPCA) empregando modelos de seleção automática de variáveis e diferentes grupos de variáveis com as informações agregadas e desagregadas das séries de inflação, tanto nacional quanto regional. Os resultados mostraram que informações com maior granularidade são capazes de adicionar maior poder preditivo porém requerem modelos com mais filtros na seleção de variáveis.

Palavras-chave: Índice de inflação, *Autometrics* e séries de tempo desagregadas

ABSTRACT

This study has the objective of evaluating different models for forecasting the national consumer's inflation index of Brazil (IPCA) using models for automatic selection of variables (Autometrics) and different groups of variables regarding aggregated and disaggregated inflation time series, both on a national level and a regional level. The results showed that more granular information are able to bring a bigger forecasting capacity, but require models with more intense filtering when selecting the variables.

Keywords: Inflation, Autometrics and disaggregated time series

Sumário

1.	Motivação.....	10
2.	Dados.....	12
3.	Análise prévia dos dados.....	14
4.	Metodologia	17
5.	Resultados	23
6.	Conclusão	32
7.	Bibliografia	33

Lista de Figuras

Figura 1 - Representação simplificada do modelo de árvore proposto por Doornik (2007).....	18
Figura 2 - Comparação entre previsões e resultados observados para inflação acumulada.....	23
Figura 3 - Erro quadrático para modelo AR(1) um passo a frente com referência em 2017(2).....	24
Figura 4 - Erro quadrático para modelo IIS um passo a frente com referência em 2017(2).....	25
Figura 5 - Erro quadrático para modelo IIS um passo a frente com referência em 2018(2).....	26
Figura 6 - Erro quadrático para grupo 0.agregado_nacional um passo a frente com referência em 2017(2).....	27
Figura 7 - Erro quadrático para grupo 5.desagregado_subgrupo_regiao um passo a frente com referência em 2017(2).....	28

Lista de Tabelas

Tabela 1 - Correlação entre índices de preços regionais e o índice de preços nacional	14
Tabela 2 - Pesos regionais por grupo de produtos	15
Tabela 3 - Correlação entre produtos por região e o índice nacional ...	15
Tabela 4 - Resultado agregados por modelo e por grupo de variáveis.	29
Tabela 5 - Resultados MCS para diferentes modelos.....	30
Tabela 6 - Resultados MCS para diferentes grupos de variáveis	31

1. Motivação

O tema central proposto para a dissertação é a comparação de diferentes abordagens para geração de projeções do Índice de Preços ao Consumidor (IPCA), principal índice de inflação brasileiro. Deseja-se estudar diferentes metodologias para modelar a série de tempo correspondente ao histórico do IPCA no Brasil, algo que pode ser feito tanto de maneira agregada quanto de maneira desagregada com uma agregação posterior dos dados desagregados. Análise do índice de maneira agregada é uma abordagem mais tradicional de séries de tempo, podendo ser empregadas diferentes modelos, como por exemplo Autoregressive integrated moving average (ARIMA), Vector autoregression (VAR) ou Markov-switching (MS), que podem oferecer maior ou menor poder de previsão quando comparados com algum benchmark, seja ele um outro modelo ou previsões de mercado.

Porém, como se trata de um índice de preços, o IPCA é resultado de uma média ponderada de uma cesta de diferentes bens nas principais regiões metropolitanas do Brasil, sendo dividido em 9 grupos, 19 subgrupos, 52 itens e 379 subitens (IBGE 2012). Portanto, por como cada subdivisão é por sua vez também uma série de tempo, é possível modelar individualmente as séries das subdivisões e depois agrega-las usando os ponderadores característicos do índice para então obter uma previsão final para o índice consolidado. Para a modelagem das séries desagregadas podem-se empregar as mesmas técnicas de modelagem de séries de tempo mencionadas para a série do índice agregado, porém espera-se que haja um maior ganho de informação através das projeções desagregadas e com sua posterior agregação do que a projeção da série agregada. Esse tipo de análise, comparando projeções agregadas versus projeções desagregadas foi feita por (CARLO 2016), na qual concluiu que um modelo desagregado com modelos desagregados da família SARIMA (Seasonal autoregressive integrated moving average) apresentava um melhor desempenho que outras técnicas, como técnicas de espaço de estados, aplicadas ao índice agregado de inflação.

Continuando desse ponto de estudo de séries desagregadas do IPCA para projeções do índice agregado, outras abordagens possíveis são os efeitos da introdução de previsões robustas na modelagem das séries de tempo desagregadas. Essas previsões robustas atuam como correções na previsão, comparando resultados de previsões de períodos anteriores contra os valores realizados da série de tempo,

reintroduzindo essas diferenças na forma de “robustificadores” da previsão. A introdução de elementos robustos na previsão de séries de tempo pode permitir a redução dos erros de previsão quando ocorrem, por exemplo, quebras estruturais na série como mudanças de patamar no índice de preço ou mudanças de tendência. Por outro lado, essas técnicas de robustificação podem introduzir muito ruído na previsão aumentando a variância das previsões quando não ocorrem tais quebras estruturais na série temporal original, fazendo com que seja importante analisar e calibrar os “robustificadores” para evitar esses efeitos secundários de aumento da variância.

Esse assunto de previsões robustas em séries de tempo foi estudado por (LARSON 2015) na modelagem de reservas e preços de petróleo nos Estados Unidos. Foi concluído que a modelagem desagregada dos preços de petróleo na presença de um robustificador de Intercept Correction (que compreende na soma do erro de previsão nas previsões seguintes) mostrou ganhos de previsão em relação à modelagem do preço agregado ou mesmo da modelagem do preço agregado utilizando informações desagregadas para a previsão, o que motiva o emprego de tais técnicas para obter ganhos de previsão na modelagem do IPCA na sua forma desagregada.

Essa dissertação tem como objetivo expandir nesses tópicos, buscando empregar outras técnicas na previsão das séries desagregadas ou procurando outras desagregações não abordadas como desagregações regionais, que possibilitem melhores previsões na modelagem do IPCA, algo que poderia ter diversas aplicações.

2. Dados

Como se trata de uma análise que busca obter ganhos de previsão utilizando as informações detalhadas das séries de inflação disponibilizadas pelo IBGE, a principal fonte de dados para esse trabalho será justamente as séries históricas de inflação. A estrutura dos dados disponibilizados relativos ao IPCA, o índice de inflação que estamos interessados em estudar, pode ser dividida basicamente em duas dimensões de detalhamento.

A primeira dimensão é o nível de produtos que corresponde as séries que acompanham os preços em uma visão menos agregada dos itens que compõe a cesta de bens que integram o índice geral, havendo quatro níveis de produto nas séries de inflação de maneira hierarquizada:

- Grupo: nível mais genérico de produto, sendo a etapa anterior ao índice geral. É composto pelos principais agrupamentos de produtos consumidos pelas famílias brasileiras, entre eles Alimentação, Habitação e Saúde. São nove os grupos que são avaliados pelo IBGE.
- Subgrupo: nível mais desagregado do nível grupo, sendo mais uma quebra no detalhamento do grupo, por exemplo o grupo Alimentação é composto por dois subgrupos: Alimentação no domicílio e Alimentação fora do domicílio. São 19 os subgrupos avaliados pelo IBGE.
- Item: Oferece um maior detalhamento do subgrupo analisado, continuando com o exemplo do grupo Alimentação, o subgrupo Alimentação no domicílio em composto por itens como Carnes, Frutas, Pescados, entre outros. São 52 os itens avaliados pelo IBGE.
- Subitem: Finalmente, o nível de subitem oferece o detalhamento mais granular dos produtos que compõe a cesta de bens do índice de inflação, são os subitens que são efetivamente medidos mensalmente pelo IBGE de modo a observar as variações de preços no território nacional. Por exemplo, no grupo de Alimentos, são subitens produtos como Arroz, Feijão, Milho, entre outros. Nesse nível mais granular são monitorados 387 subitens pelo IBGE.

A lista completa dos grupos, subgrupos, itens e subitens avaliados pelo IBGE encontra-se no anexo.

Os dados desagregados são utilizados então para compor a versão agregada do índice de preço, utilizando pesos que correspondem ao consumo das famílias de determinado produto, conforme descrito no manual Método de Cálculos, disponibilizado pelo Sistema Nacional de Índices de Preços ao Consumidor. (IBGE 2012)

Adicionalmente aos dados desagregados em diferentes níveis hierárquicos, o IBGE disponibiliza mais uma dimensão, a dimensão de região. A pesquisa de preços realizada para a composição do IPCA ocorre em diferentes regiões metropolitanas brasileiras. Cada subitem, conforme explicado anteriormente, é levantado em cada região metropolitana e depois consolidado em nível nacional. O peso utilizado para compor o índice nacional segue a proporção da população urbana para cada estado, tendo como fonte a Pesquisa de Orçamentos Familiares (POF) 2008-2009, como também descrito no manual Método de Cálculos, disponibilizado pelo Sistema Nacional de Índices de Preços ao Consumidor.

Desde 2012, data da última mudança na metodologia de cálculo do IPCA pelo IBGE, são analisadas 10 regiões metropolitanas no Brasil, que posteriormente irão compor o índice nacional.

3. Análise prévia dos dados

A principal hipótese a ser analisada nesse estudo é que a incorporação de informações das séries desagregadas (tanto por produto quanto por região), ofereceriam ganhos de previsão em relação as projeções que utilizem somente informações da série nacional agregada. Antes de se realizar regressões às séries desagregadas e aplicar a metodologia que será descrita a seguir, vale a pena observar como esses dados desagregados se relacionam com índice agregado e como estes variam no tempo.

Observando como as diferentes regiões se correlacionam com o índice geral, pode-se ver que existe alta correlação entre o índice nacional e as regiões com maior adensamento urbano (como São Paulo e Belo Horizonte), como pode ser visto na Tabela 1. Porém a ausência de uma correlação perfeita mostra também que utilizando informações mais granulares a nível de região pode ser possível extrair mais informações para previsões, assim possivelmente obtendo-se ganhos de previsão com a introdução dessas variáveis, não apenas utilizando o índice nacional.

Região	Correlação (%)
Belem_PA	70,0%
BeloHorizonte_MG	88,7%
Curitiba_PR	82,0%
Fortaleza_CE	80,0%
PortoAlegre_RS	87,0%
Recife_PE	72,1%
RiodeJaneiro_RJ	84,2%
Salvador_BA	78,8%
SaoPaulo_SP	91,9%

Tabela 1 - Correlação entre índices de preços regionais e o índice de preços nacional

Outra análise prévia que pode ser feita é com relação aos pesos entre grupos e subgrupos entre as diferentes regiões brasileiras monitoradas. Como se ver na Tabela 2, a distribuição de pesos de produtos entre as regiões não é homogênea. Embora a ordenação de pesos entre grupos seja similar entre as regiões, o peso relativo de cada grupo na cesta de preços locais varia de região para região.

Isso indica que poderiam se haver ganhos ao se projetar cada índice de preços por grupo para cada região para então compor o índice geral de preços por região para finalmente agrupá-los e obter previsões para o índice geral nacional.

Grupo	Região								
	Belem_PA	BeloHorizonte_MG	Curitiba_PR	Fortaleza_CE	PortoAlegre_RS	Recife_PE	RiodeJaneiro_RJ	Salvador_BA	SaoPaulo_SP
1.Alimentacao e bebidas	34,2	22,0	23,6	32,0	25,8	27,0	23,6	27,9	23,4
2.Habitacao	12,5	15,8	16,3	14,0	14,0	14,0	17,0	14,3	14,6
3.Artigos de residencia	5,2	5,1	4,4	4,4	4,7	5,0	3,6	4,6	3,8
4.Vestuario	8,6	6,6	7,2	7,1	6,5	7,6	4,8	6,9	5,6
5.Transportes	12,7	18,5	20,0	16,0	18,4	15,2	17,7	19,2	20,2
6.Saude e cuidados pessoais	10,5	11,0	11,5	9,8	11,3	12,8	12,1	10,4	12,0
7.Despesas pessoais	8,2	12,0	9,9	9,1	10,5	9,9	11,3	8,3	11,2
8.Educacao	4,6	4,7	3,4	4,3	4,5	4,6	5,1	4,4	5,0
9.Comunicacao	3,4	4,2	3,7	3,3	4,0	3,8	4,9	4,1	4,1

Tabela 2 - Pesos regionais por grupo de produtos

Finalmente, pode-se observar como cada série de preços por grupo de cada região monitorada se relaciona com o a série nacional do mesmo grupo de produtos. Isso revela maior variância entre as correlações pois além das variações de pesos das regiões no índice nacional e das variações entre os pesos de cada grupo na cesta de cada região existem comportamentos regionais na determinação de cada preço dos grupos de produtos.

Região	Grupo								
	1.Alimentacao e bebidas	2.Habitacao	3.Artigos de residencia	4.Vestuario	5.Transportes	6.Saude e cuidados pessoais	7.Despesas pessoais	8.Educacao	9.Comunicacao
Belem_PA	79,8%	49,2%	53,6%	51,5%	45,4%	76,5%	73,0%	98,4%	76,7%
BeloHorizonte_MG	91,4%	85,2%	50,2%	66,4%	83,3%	81,7%	71,8%	98,7%	80,7%
Curitiba_PR	85,9%	84,7%	59,7%	64,0%	72,5%	75,0%	76,1%	98,8%	88,7%
Fortaleza_CE	84,8%	76,6%	37,1%	62,2%	58,7%	75,5%	73,0%	21,2%	55,7%
PortoAlegre_RS	80,7%	83,0%	51,3%	72,9%	68,3%	88,2%	82,3%	85,2%	91,4%
Recife_PE	82,9%	55,8%	51,8%	64,9%	59,5%	78,9%	55,1%	94,6%	59,0%
RiodeJaneiro_RJ	85,1%	77,3%	64,1%	69,4%	69,9%	80,7%	62,2%	98,7%	88,4%
Salvador_BA	82,3%	52,5%	34,4%	50,8%	64,6%	76,5%	47,9%	97,7%	82,8%
SaoPaulo_SP	90,1%	90,5%	83,8%	86,8%	93,8%	93,4%	79,6%	98,5%	82,0%

Tabela 3 - Correlação entre produtos por região e o índice nacional (cores indicam maior correlação no caso de verde e menor correlação no caso de vermelho)

A Tabela 3 mostra como cada série de preços de cada grupo por região se correlaciona com o índice de preços nacional desse determinado grupo de produtos, as cores utilizadas variam numa escala de verde (correlações mais próximas de 100%) e vermelho (correlações mais próximas de 0%) quanto a correlação das variações desse índice local com o índice nacional agregado. Isso mostra que a variação de preços por grupo não é uniforme no território nacional e que a introdução dessa dimensão de grupo por região pode agregar valor às projeções do índice geral nacional.

4. Metodologia

O estudo em questão tem como objetivo obter ganhos de previsão no índice geral de preços através da introdução de novas variáveis que contém informações mais granulares relativas aos índices de preços regionais e por grupo. É necessário estabelecer um *benchmark* para comparação dos ganhos de previsão. Como hipótese, tem-se que a introdução de dados mais granulares sobre a inflação conterá mais informação do que apenas o índice geral agregado, portanto deseja-se comparar o poder preditivo de modelagens que envolvam apenas esse índice agregado, sem nenhuma outra informação.

Uma vez definido um grupo básico sendo esse o índice agregado nacional, serão analisados quatro outros grupos de variáveis com as quais espera-se obter um maior ganho de informação. Eles são (1) variáveis nacionais desagregadas a nível grupo, (2) variáveis nacionais desagregadas a nível subgrupo, (3) variáveis regionais agregadas, (4) variáveis regionais desagregadas a nível grupo e (5) variáveis regionais desagregadas a nível subgrupo.

Para cada grupo de variáveis, cada série de tempo dos índices de preços em sua determinada granularidade é desenvolvido um modelo de projeção conforme os modelos que serão discutidos a seguir e em seguida reagregados utilizando os respectivos pesos na cesta de bens de cada região e depois no peso de cada região no índice nacional. No caso dos índices nacionais, o peso de cada item já corresponde ao peso final. Cada modelo foi avaliado utilizando variáveis que vão do início da série que tem início em no ano de 1999.

Para a modelagem geral, todas as séries de tempo desagregadas seguem o mesmo modelo autoregressivo com doze defasagens, de maneira geral:

$$y_{i_t} = \alpha_i + \beta_{i_1}y_{i_{t-1}} + \beta_{i_2}y_{i_{t-2}} + \dots + \beta_{i_{12}}y_{i_{t-12}}$$

Para cada uma das variáveis desagregadas.

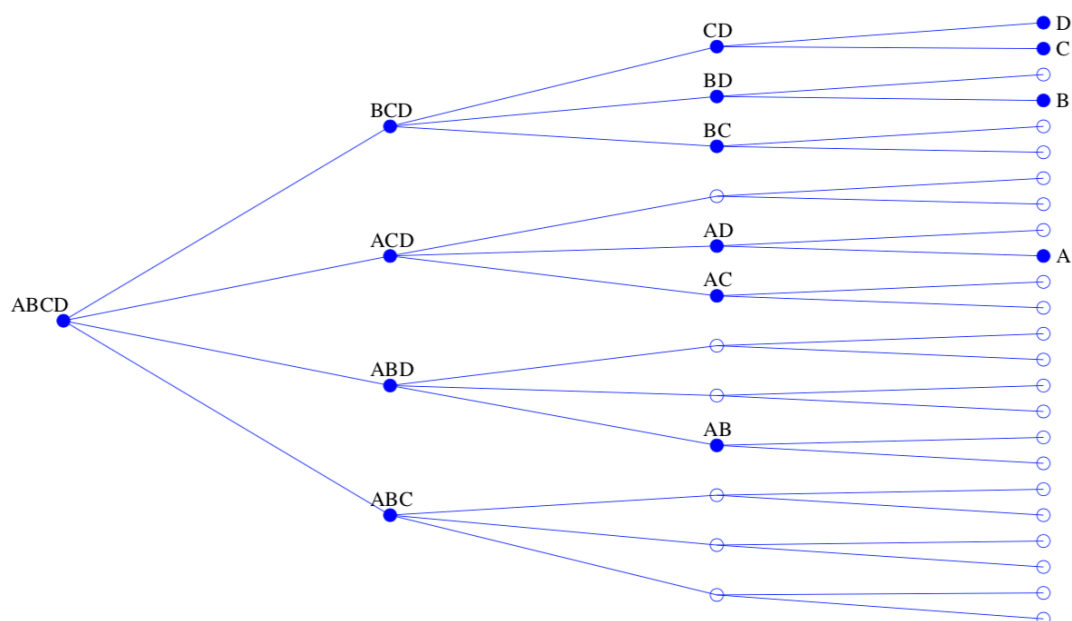
Uma maneira de lidar com tantas variáveis e defasagens envolvidas em cada uma das regressões é a utilização de um algoritmo de seleção automática de variáveis “*Autometrics*” conforme desenvolvido por Hendry e Doornik (2014).

Inicialmente proposto por Doornik (2007), o algoritmo “*Autometrics*” tem como ponto central um modelo contendo todas as variáveis possíveis, o chamado modelo geral

irrestrito (ou “*General unrestricted model*” – GUM, no original em inglês), que contem informações relevantes sobre o processo que está sendo avaliado e que seja estatisticamente bem-comportado. No caso em questão, seriam os modelos desagregados com as múltiplas defasagens.

A partir do GUM, podem ser definido todo um subespaço de modelos que contém menos variáveis que o modelo geral, eliminando uma variável por vez e recalculando o modelo sem essa variável eliminada.

Figura 1 - Representação simplificada do modelo de árvore proposto por Doornik (2007)



A Figura 1 mostra uma representação da definição desse subespaço de variáveis a partir do GUM como uma árvore na qual cada ponto representa um conjunto de variáveis diferentes e quanto mais a esquerda o ponto se encontra, menos variáveis ele possui, sendo que a cada ramificação de um ponto isso representa o processo de eliminação de uma variável do modelo anterior. Cada modelo, como definido nessa árvore de modelos é avaliado tanto rodando um “*backtest*” com relação ao modelo geral quanto rodando um teste de diagnóstico. Caso um modelo falhe em qualquer um dos testes, esse modelo é julgado como não significativo.

Como varrer todos os nós dessa árvore de modelos pode ser altamente caro computacionalmente e ineficiente na busca pelo melhor conjunto de variáveis, Doornik (2007) propõe algumas estratégias para otimizar essa busca:

- “*pruning*” (ou poda): quando um modelo (ou um nó) é definido como não significativo segundo qualquer critério, todos os nós que se encontram a direita desse nó, ou seja, que tem modelos derivados desse modelo não significativo, são eliminados automaticamente da busca.
- “*Bunching*” (ou agrupamento): ao invés de tentar eliminar apenas uma variável por vez, o algoritmo tenta eliminar um grupo de variáveis agrupadas, testando a significância desse grupo de variáveis. Caso o modelo julgue esse agrupamento de variáveis como não significantes, então é possível pular diversos ramos na árvore de procura, eliminando de uma vez só alguns nós e ramos, tornando a busca mais efetiva.
- “*Chopping*” (ou corte): quando um agrupamento de variáveis, como definido anteriormente, é julgado como pouco significativo, todos os ramos decorrentes desses modelos podem ser eliminados da árvore original, limitando ainda mais a busca na seleção de variáveis.

Nas regressões avaliadas na seleção automática de variáveis foram empregados os seguintes métodos de avaliação dentro do algoritmo “*Autometrics*”, uma vez que os algoritmos desenvolvidos permitem não só a seleção de variáveis com base em sua significância estatística, como também alguns tratamentos de *outliers* dentro da série de tempo:

- “*No*”: Nenhuma seleção feita
- “*None*”: Nenhum teste para detecção de anomalias adicionado, só seleção

Alguns dos tratamentos de *outliers* avaliados e empregados dentro do algoritmo “*Autometrics*” foram:

- “*Impulse-indicator saturation*” (IIS): Teste para detecção de saturação na presença de *outliers*, sendo que esses *outliers* podem acontecer a qualquer momento, no qual cada observação é tratada como um impulso, podendo eliminar esses impulsos que não sejam significativos

- "*Dummy Impulse-indicator saturation*" (DIIS): Teste para detecção de saturação na presença de *outliers* porém com a utilização de variáveis *dummy*, no qual cada observação é transformada numa variável *dummy* e caso ela seja significativa, essa variável é incorporada ao modelo final.
- "*Step-indicator saturation*" (SIS): Teste para detecção de saturação na presença de *outliers* para um intervalo de tempo (não apenas para um instante como no caso dos impulsos). Esse teste é utilizado para capturar o efeito de quebras estruturais que sejam persistentes no tempo, por mais de um período, isso ajuda a caracterizar mudanças dentro do cenário analisado no tempo e permite captura-las de maneira mais eficiente do que usando variáveis *dummy* para cada instante após a mudança estrutural, como aconteceria no caso do teste DIIS.

Os testes do "*Autometrics*" podem ser combinados entre si para fornecer modelos de avaliação mais complexos. Sendo assim, os modelos escolhidos para avaliação dos grupos de variáveis são:

- No
- None
- IIS
- DIIS
- IIS + DIIS
- IIS + DIIS + SIS

Os modelos selecionados pelo algoritmo "*Autometrics*" serão comparados com um modelo mais simples, no caso foi escolhida uma regressão simples AR(1), ou seja uma regressão autoregressiva com lag de 1 período para ser uma regressão *benchmark* das demais.

Finalmente, para a avaliação do desempenho de cada modelo e grupo de variáveis utilizadas para prever o índice de inflação nacional é preciso escolher uma função para calcular o erro das previsões em relação ao índice de fato observado. A função escolhida foi o erro logarítmico dos índices:

$$erro_t = \ln Previsto_t - \ln Observado_t$$

Para evitar desvios negativos, é mais indicado utilizar o erro quadrático do erro logarítmico calculado anteriormente:

$$erro_t^2 = (erro_t)^2$$

Uma maneira para comparar simultaneamente as diferentes combinações de modelos e grupos de variáveis é utilizar o método de “*Model Confidence Set*” (MCS) desenvolvido por Hansen (2014) e consiste em testes em sequência para a construção de um conjunto de modelos superior, onde a cada teste é verificada a hipótese nula de que dado um grau de confiança os modelos possuem o mesmo poder preditivo. O MCS tem como insumo alguma função de perda para representar o erro do modelo, para o caso foi empregada a mesma função de erro logarítmica que foi utilizada para calcular o erro quadrático médio anteriormente.

O MCS é definido como o conjunto de modelos que possuem estatisticamente o mesmo erro como definido pela função de perda anteriormente definida. A hipótese nula que é testada em cada etapa é da forma:

$$H_0: E(d_{ij,t}) = 0$$

Onde i e j são modelos dentro do conjunto de possíveis modelos e a variável d (que mede a diferença de performance entre os conjuntos) é definida como:

$$d_{ij,t} = L_{i,t} - L_{j,t}$$

Onde L é a função de perda em um instante t.

A partir dessa hipótese nula, o MCS atua de maneira iterativa seguindo uma regra de eliminação. No início, todos os modelos são incluídos no conjunto de modelos superiores. Em seguida, é testada a hipótese nula anteriormente definida para cada um dos pares de modelos presentes, mediante um grau de confiança estatístico.

Caso, o teste seja aceito para todos pares testados, o conjunto inicial já é definido com o conjunto superior, caso contrário, é seguida alguma regra de eliminação, excluindo os modelos com maior diferença de performance e testando a hipótese nula novamente. Esse processo de teste e eliminação é repetido até que a hipótese nula seja aceita para todos os pares de modelos do conjunto, sendo esse conjunto finalmente definido como o conjunto superior.

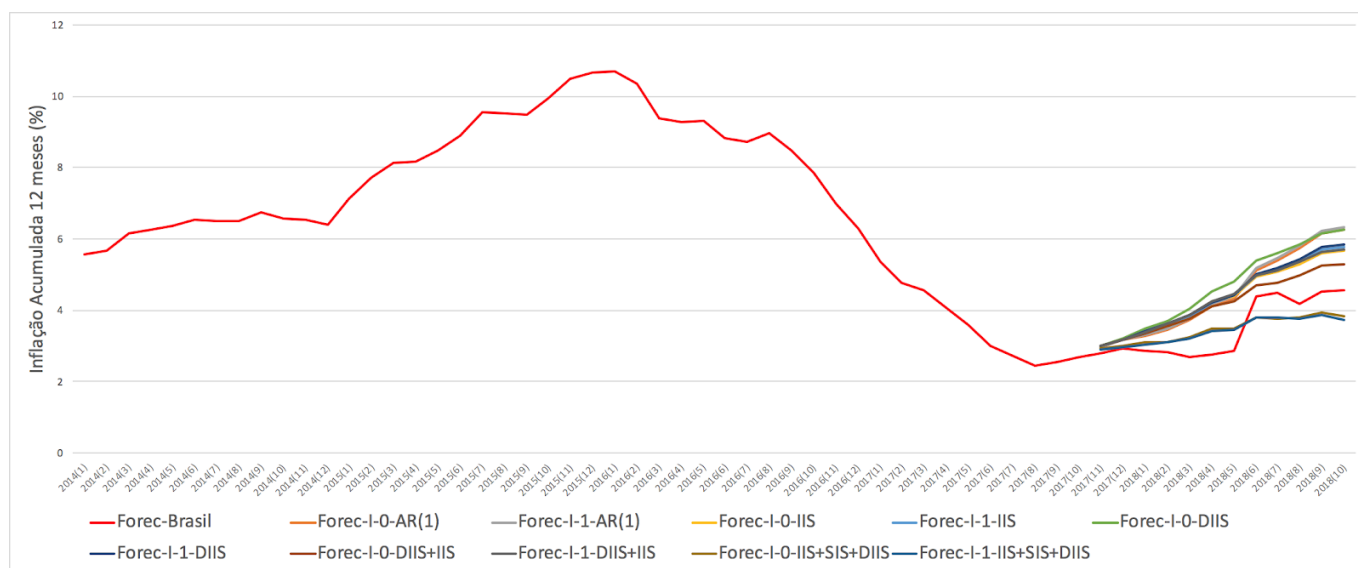
5. Resultados

Conforme descrito na seção de Metodologia, serão comparadas as previsões feitas a partir de dados dos diversos grupos com diferentes níveis de informação desagregada para recompor e reagregar o índice geral de preços nacional. Cada série regional foi modelada individualmente e depois utilizando a composição de pesos de cada região o somatório dos índices regionais previstos compõe a previsão do índice nacional.

Foram calculadas previsões até 36 meses a frente do mês de referência utilizado em cada uma das regressões. As previsões foram feitas iniciando no mês de janeiro/19 e voltando 24 meses produzindo assim 24 séries de previsões para cada uma das combinações de modelos de seleção de variáveis e grupo de variáveis de inflação. Assim, o número total de séries calculadas foi de:

24 meses de previsões x 6 grupos de variáveis x 7 modelos = 1.008 séries de previsões

Figura 2 - Comparação entre previsões e resultados observados para inflação acumulada



A Figura 2 mostra as previsões realizadas por cada modelo para os 12 meses seguintes, comparando o resultado das previsões com a série realizada da inflação no período.

Como uma forma de avaliação é possível observar e comparar por exemplo os erros quadráticos médios observados nos períodos para cada combinação de modelo empregado e grupo de variáveis.

Para visualizar como esses erros evoluem no tempo uma alternativa é fixar um mês de referência a partir do qual as projeções serão feitas, fixando também um modelo específico e avaliando como a performance de cada grupo de variáveis em cada um dos meses subsequentes em relação ao observado, calculando-se o erro conforme a função de erro logarítmico definida na seção de Metodologia anteriormente. A primeira análise, mostrada na Figura 3, pode ser feita com modelo definido como referência aos demais modelos que empregam o algoritmo do “*Autometrics*”, o modelo autorregressivo AR(1) e com mês de referência mais antigo, ou seja com o maior número de comparações possíveis com os dados observados, que no caso é fevereiro/17.

Figura 3 - Erro quadrático para modelo AR(1) um passo a frente com referência em 2017(2)

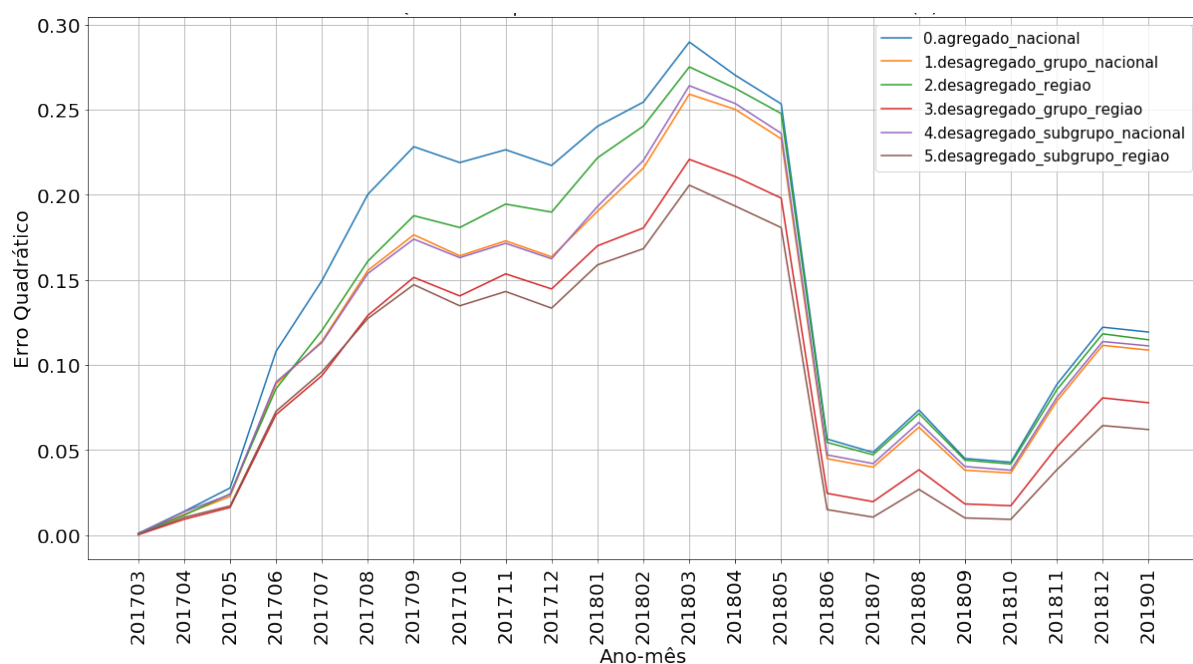


A figura 3 mostra a evolução do erro quadrático a cada mês para os diferentes grupos de variáveis, desde o mais amplo (0.agregado_nacional, que é o índice nacional agregado) até o mais granular (5.desagregado_subgrupo_regiao, que corresponde aos índices regionais quebrados em subgrupos no nível produto). Esse erro foi definido com o erro do índice acumulado em 12 meses medido um passo a frente para cada um dos modelos. Como pode-se ver, para um modelo simples como o AR(1), o acréscimo de mais variáveis sem tratamentos ou filtros ao modelo acabam produzindo um resultado muito próximo ou mesmo inferior ao resultados obtidos pelos modelos com mais variáveis disponíveis, julgando pela proximidade dos erros obtidos

mês a mês. Vale a pena ressaltar o formato dessa curva de erro, como se trata de erros um passo a frente, seria esperado que esse erro crescesse a partir do mês de referência, o que de fato ocorre nos primeiros meses. No entanto, no mês 2018/05 existe uma queda brusca nesse erro em relação ao índice acumulado. Isso pode ser explicado pelo fato de que nos meses anteriores a 2017, a inflação vinha em níveis elevados, fazendo com que os modelos tivessem um erro positivo em relação aos índices observados, o que se manteve durante 2017 e o início de 2018. Porém, em 2018/05 houve a greve dos caminhoneiros no Brasil, afetando os índices de preço e colocando o índice acumulado em patamares mais elevados que anteriormente observados e esse deslocamento positivo do índice fez com o que o erro das previsões caísse, uma vez que o previsto se encontrava maior que o observado justificando a queda supostamente anormal das curvas de erro. Esse formato de curva de erro crescente com uma queda em 2018/05 é algo que irá se repetir para diferentes gráficos e análises a seguir.

Utilizando agora um modelo com uma seleção de variáveis empregando o “Autometrics”, por exemplo, um modelo “*Impulse-indicator saturation*” (IIS) para os mesmos grupos de variáveis, obtemos o seguinte resultado.

Figura 4 - Erro quadrático para modelo IIS um passo a frente com referência em 2017(2)

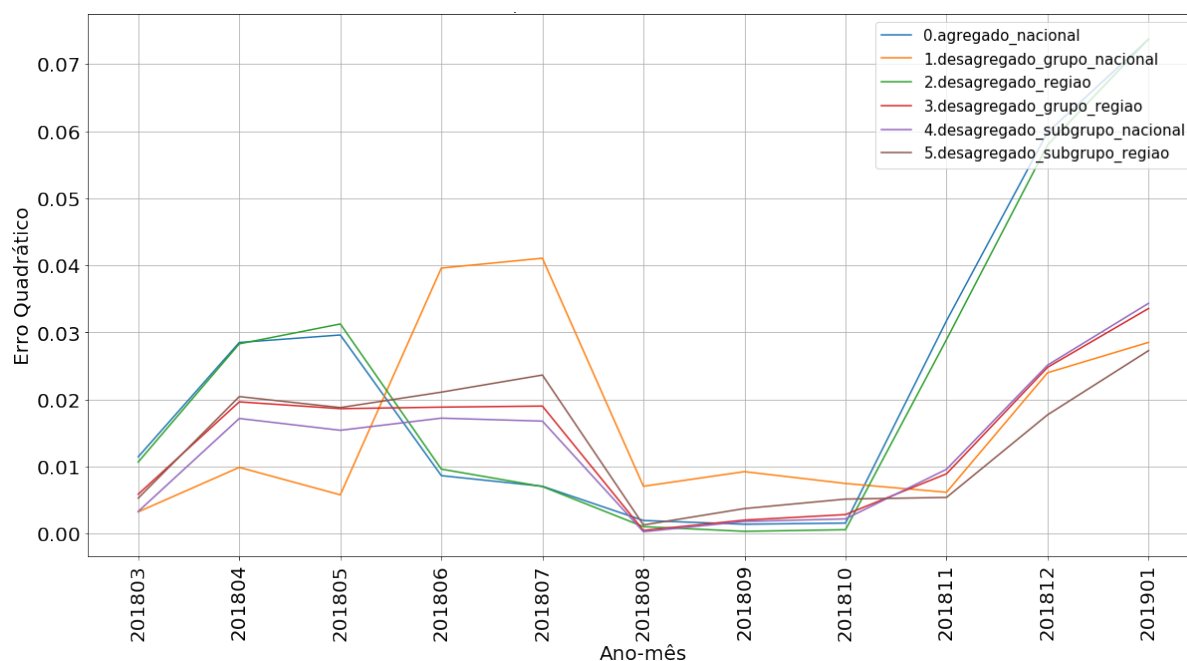


Com um modelo que emprega uma seleção de variáveis antes de inclui-las no modelo final de previsão é possível ver que:

- Foram obtidos melhores resultados de previsão para todos os conjuntos de variáveis, como pode-se ver pelos menores erros em relação ao modelo AR(1)
- Esses ganhos de previsão foram mais sensíveis nos grupos de variáveis com maior granularidade de informação, sendo que nesse modelo todos os conjuntos de variáveis obtêm melhores resultados que o do índice agregado nacional. Vale também ressaltar que os dois grupos com menor erro durante todo o período foram os grupos de variáveis que contem informação na granularidade de região (3.desagregado_grupo_regiao e 5.desagregado_subgrupo_regiao).

Utilizando outras janelas de tempo para avaliar os diferentes grupos de variáveis para um dado modelo de previsão com outros meses de referência é possível ver a consistência dos resultados no tempo.

Figura 5 - Erro quadrático para modelo IIS um passo a frente com referência em 2018(2)

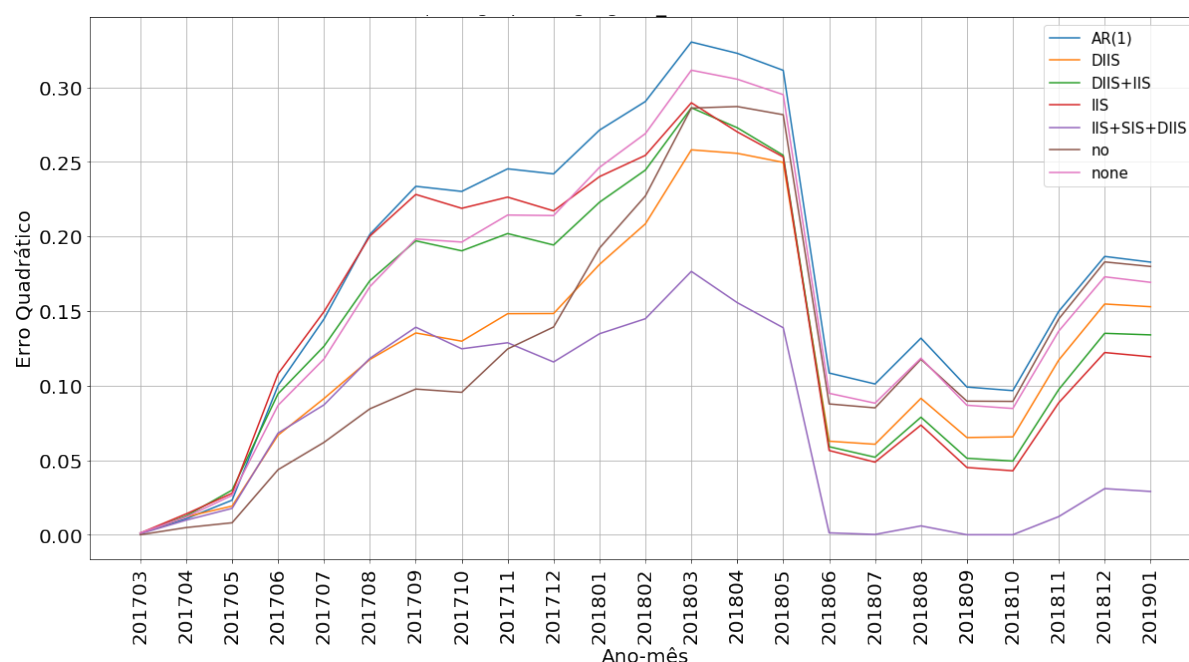


Utilizando agora fevereiro/18 ao invés de fevereiro/17 para o modelo IIS os resultados para os grupos de variáveis mais desagregados se mantem comparando 12 meses ao invés de 24 meses, como no gráfico anterior.

É possível também inverter a referência da análise anterior e fixar não o modelo empregado para as previsões, mas sim os grupos de variáveis, plotando os gráficos que mostram a evolução do erro no tempo a partir de uma referência.

Da mesma maneira que o modelo AR(1) se apresenta como uma referência para os modelos, o grupo de variáveis com apenas o índice nacional agregado apresenta-se como uma referência para a performance dos grupos de variáveis. Sendo assim, a primeira abordagem para analisar os modelos dado um grupo de variáveis é observando os erros cometidos para os modelos empregando o índice nacional agregado.

Figura 6 - Erro quadrático para grupo 0.agregado_nacional um passo a frente com referência em 2017(2)

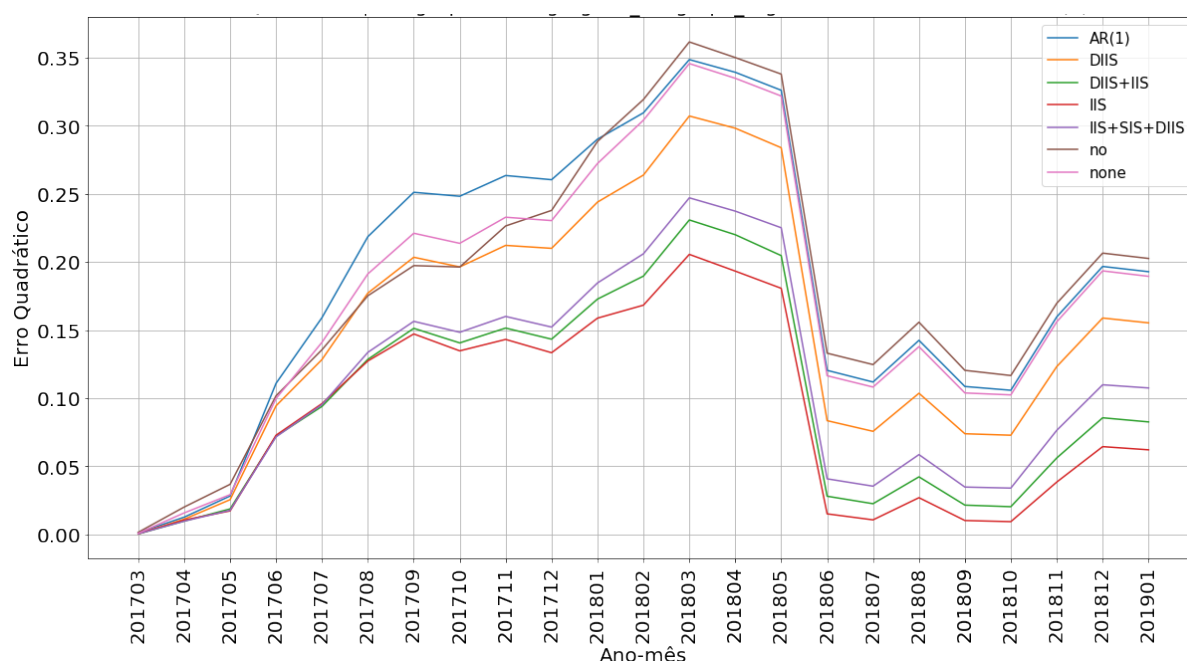


Da mesma maneira que para a abordagem anterior, foi utilizado o mês de referência de fevereiro/17 e o cálculo dos erros mês a mês do índice de inflação acumulado em 12 meses. Para o menor número de variáveis possível (apenas o índice), os modelos que empregam seleções e filtros mais elaborados como "Impulse-indicator saturation" (IIS), "Dummy Impulse-indicator saturation" (DIIS) e "Step-indicator saturation" (SIS) combinados ao mesmo tempo apresentam melhores resultados que os modelos mais simples, como por exemplo o "benchmark" AR(1). Como notado anteriormente, é

possível ver um comportamento supostamente anômalo em 2018/05 mas que pode ser explicado pela elevação dos níveis de inflação pela greve dos caminhoneiros.

Por outro lado, o grupo de variáveis com o maior número de variáveis disponível é o desagregado regional com nível de subgrupo. Nesse conjunto de variáveis são empregadas 108 variáveis de índices de subgrupos de produtos em cada uma das regiões mapeadas.

Figura 7 - Erro quadrático para grupo 5.desagregado_subgrupo_regiao um passo a frente com referência em 2017(2)



Novamente nesse comparativo entre os diferentes modelos de previsão é possível verificar que os modelos com maior controle das variáveis introduzidas possuem uma performance melhor que os modelos sem esse tipo de filtros. Vale também notar que para o mesmo modelo, o acréscimo de variáveis na regressão possui um efeito positivo na redução dos erros.

Compilando todos os resultados na forma de uma tabela permite a comparação simultânea de todas as combinações de modelos e grupo de variáveis. A tabela foi calculada utilizando também como referência o mês de fevereiro/17, consolidando os erros quadráticos médios de cada período de previsão.

model	0.agregado_nacional	1.desagregado_grupo_nacional	2.desagregado_regiao	3.desagregado_grupo_regiao	4.desagregado_subgrupo_nacional	5.desagregado_subgrupo_regiao
AR(1)	0.17457	0.16731	0.18513	0.18143	0.17507	0.18723
DIIS	0.12145	0.11356	0.17041	0.14104	0.16199	0.15230
DIIS+IIS	0.13724	0.11258	0.14546	0.10385	0.11080	0.09942
IIS	0.14329	0.11919	0.12965	0.09646	0.12056	0.08811
IIS+SIS+DIIS	0.07132	0.13475	0.06926	0.06134	0.12705	0.11055
no	0.12658	0.16273	0.14139	0.16978	0.17216	0.18329
none	0.15707	0.16260	0.16958	0.16768	0.16689	0.17663

Tabela 4 - Resultado agregados por modelo e por grupo de variáveis

Como descrito na seção de Metodologia, o MCS pode ser empregado para comparar os diferentes modelos. As simulações empregando o MCS foram divididas em dois principais grupos:

- Testes empregando todos as combinações de modelos e grupo de variáveis ao mesmo tempo
- Testes isolando um determinado modelo de previsão e comparando os diferentes grupos de variáveis para avaliar o efeito isolado da adição de mais variáveis para ganhos preditivos
- Testes isolando um determinado grupo de variáveis e comparando os diferentes modelos para avaliar o efeito isolado dos modelos para ganhos preditivos

Para o primeiro grupo de testes do MCS foram utilizadas todas as variáveis para diferentes níveis de confiança (5%, 10% e 25%). Quando todas as variáveis foram colocadas no teste, mesmo colocando diferentes níveis de confiança, ainda foram selecionadas muitas variáveis como o grupo superior de modelos. Sendo assim, é mais interessante isolar e controlar os grupos de variáveis e os modelos, primeiro fixando os modelos e variando os grupos de variáveis e depois fixando os grupos de variáveis e encontrando os melhores modelos para esse grupo de variáveis.

Em seguida, para o segundo grupo de testes do MCS foram realizados sete testes, um para cada modelo estimado, utilizando dentro desses testes as seis previsões com as diferentes granularidades de variáveis. Todos os testes desse segundo grupo foram feitos utilizando 10% de nível de confiança.

model	Melhor grupo de variáveis
AR(1)	1.desagregado_grupo_nacional
DIIS	0.agregado_nacional / 1.desagregado_grupo_nacional
DIIS+IIS	5.desagregado_subgrupo_regiao
IIS	5.desagregado_subgrupo_regiao
IIS+SIS+DIIS	3.desagregado_grupo_regiao
no	0.agregado_nacional
none	0.agregado_nacional

Tabela 5 - Resultados MCS para diferentes modelos

Como esperado, os modelos com filtros mais elaborados foram os que mais se beneficiaram pelo acréscimo de variáveis mais granulares, sendo que para esses os grupos de variáveis escolhidos foram os grupos com maior número de séries de preços como as quebradas por subgrupo e por região.

Finalmente, para o terceiro grupo de testes do MCS foram realizados seis testes, um para cada grupo de variáveis estimado, utilizando dentro desses testes os sete modelos de previsões. Todos os testes desse segundo grupo foram feitos utilizando 10% de nível de confiança.

Grupo de variáveis	Melhor modelo
0.agregado_nacional	*Todos os modelos selecionados como superiores
1.desagregado_grupo_nacional	ISS
2.desagregado_regiao	ISS / ISS+DIIS / DIIS+IIS+DIIS
3.desagregado_grupo_regiao	ISS+DIIS
4.desagregado_subgrupo_nacional	ISS+DIIS
5.desagregado_subgrupo_regiao	ISS+DIIS

Tabela 6 - Resultados MCS para diferentes grupos de variáveis

Também é possível ver na seleção dos melhores modelos que os melhores modelos para grupos de variáveis com maior granularidade foram os modelos com maiores filtros na seleção de variáveis do “*Autometrics*”.

6. Conclusão

O principal objetivo desse estudo era avaliar se a introdução de introduções mais granulares relativos a índices de inflação em nível de produto e em nível de região seria capaz de obter maiores ganhos de previsão. Foram analisados os dados a nível produto indo desde níveis agregados até quebrados por subgrupos, no total de 19 itens, e analisando os dados a nível região, comparando o índice nacional agregado e os índices separados por região metropolitana que são ponderados para compor o índice nacional. Utilizando os algoritmos de seleção automática de variáveis “Autometrics” para a construção de modelos de previsão das séries de inflação, foi possível encontrar ganhos de previsão pela introdução de uma maior granularidade de variáveis nos modelos. Esses ganhos com variáveis mais detalhadas foram maiores para os modelos que empregam filtros de seleção mais robustos na seleção de variáveis, como por exemplo o modelo "Impulse-indicator saturation" (IIS), o modelo "Dummy Impulse-indicator saturation" (DIIS) ou uma combinação desses modelos. Para modelos mais simples como um modelo autorregressivo AR(1) no entrando, a introdução de mais informações sem um tratamento na seleção dessas informações não foi muito efetivo, o que se mostrou pelo fato desses modelos terem como melhor grupo de variáveis grupos de variáveis mais condensadas e menos granulares, algo que pode ser explicado pelos ruídos que talvez essas informações de produto ou região criem nos modelos simples como o AR(1), piorando sua performance.

7. Bibliografia

Carlo, T. C.; Marçal, E. F. *Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon*. Applied Economics, 48:50, 2016

IBGE, Sistema Nacional de Índices de Preços ao Consumidor, 2012. Disponível em: https://ww2.ibge.gov.br/home/estatistica/indicadores/precos/inpc_ipca/defaultinpc.sh
[tm](#)

Larson, W., *Forecasting an Aggregate in the Presence of Structural Breaks in the Disaggregates*. Research Program on Forecasting, 2015

Dees, S., et. Al; *Exploring the international linkages of the euro area: a global var analysis*. Journal of Applied Econometrics, 2007

Barbosa, B., Marçal, E. F; *Modeling How Macroeconomic Shocks Affect Regional Employment: Analyzing the Brazilian Formal Labor Market using the Global VAR Approach*. CEMAP Working Papers, 2018

Korobilis, D.; *Machine Learning Macroeconometrics: A Primer*, Essex Finance Centre – Working Papers, 2018

Silva, A., Marçal, E. F; *Descobrendo Modelos de Previsão para a Inflação Brasileira: Uma Análise a partir de uma Gama Ampla de Indicadores*. Estudos Econômicos, 2018

Doornik, J.A.; *Autometrics*, Department of Economics, University of Oxford - 2007

Doornik, J.A., Hendry, D.F.; *Statistical model selection with 'big data'*, Institute for New Economic Thinking, Oxford Martin School, and Economics Department, Oxford University, UK. - Discussion Paper Series, 2014

Hansen, P.R., Lunde., A, Nason, J.M.; *The model confidence set*. Econometrica, 79(2), 453-497, 2011