



Universidade do Estado do Rio de Janeiro

Centro de Ciências Sociais

Faculdade de Ciências Econômicas

Carlos Henrique Dias Cordeiro de Castro

**Previsão da série temporal da inflação usando modelos orientados por score
dinâmico e métodos de aprimoramento de acurácia preditiva: o caso do
Brasil**

Rio de Janeiro

2023

Carlos Henrique Dias Cordeiro de Castro

Previsão da série temporal da inflação usando modelos orientados por score dinâmico e métodos de aprimoramento de acurácia preditiva: o caso do Brasil



Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Ciências Econômicas, da Universidade do Estado do Rio de Janeiro. Área de concentração: Métodos Quantitativos.

Orientador: Prof. Dr. Fernando Antônio Lucena Aiube

Rio de Janeiro

2023

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CCS/B

C355 Castro, Carlos Henrique Dias Cordeiro de.
Previsão da série temporal da inflação usando modelos orientados por score dinâmico e métodos de aprimoramento de acurácia preditiva: o caso do Brasil / Carlos Henrique Dias Cordeiro de Castro. – 2023.
93 f.

Orientador: Prof. Dr. Fernando Antônio Lucena Aiube.
Tese (doutorado) – Universidade do Estado do Rio de Janeiro, Faculdade de Ciências Econômicas.
Bibliografia: f. 86-92.

1. Inflação – Brasil – Teses. 2. Previsão econômica – Brasil – Teses. 3. Modelos econométricos – Teses. 4. Economia – Brasil – Teses I. Aiube, Fernando Antônio Lucena. II. Universidade do Estado do Rio de Janeiro. Faculdade de Ciências Econômicas. III. Título.

CDU 336.2(81): 330.43

Bibliotecário: Fabiano Salgueiro CRB7/6974

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Assinatura

Data

Carlos Henrique Dias Cordeiro de Castro

Previsão da série temporal da inflação usando modelos orientados por score dinâmico e métodos de aprimoramento de acurácia preditiva: o caso do Brasil

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação da Faculdade de Ciências Econômicas, da Universidade do Estado do Rio de Janeiro. Área de concentração: Métodos Quantitativos.

Aprovada em 13 de fevereiro de 2023.

Banca Examinadora:

Prof. Dr. Fernando Antonio Lucena Aiube (Orientador)
Faculdade de Ciências Econômicas - UERJ

Prof. Dr. Andrea Ugolini
Faculdade de Ciências Econômicas - UERJ

Prof. Dr. Edison Americo Huarsaya Tito
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Dr. Valdemar Rodrigues de Pinho Neto
EPGE Escola Brasileira de Economia e Finanças

Prof. Dr. Ariel Levy
Universidade Federal Fluminense

Rio de Janeiro
2023

AGRADECIMENTOS

Agradeço ao professor Fernando Aiube por toda a orientação na realização deste trabalho. Também agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) (Código de Financiamento 001) pelo apoio financeiro para o presente trabalho. Por fim, agradeço aos meus pais e familiares por todo o apoio durante meus anos de estudo.

RESUMO

CASTRO, Carlos H. D. C. de. *Previsão da série temporal da inflação usando modelos orientados por score dinâmico e métodos de aprimoramento de acurácia preditiva: o caso do Brasil*, 2023. 93 f. Tese (Doutorado em Ciências Econômicas) - Faculdade de Ciências Econômicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

Esta tese analisa as previsões mensais da inflação brasileira, medida pelo IPCA, obtidas por modelos baseados na abordagem GAS (Generalized Autoregressive Score) desenvolvido por Creal et al. (2008) e Harvey e Chakravarty (2008) com as previsões divulgadas pelo Boletim Focus, além de outros modelos competidores comumente usados na literatura. Os modelos GAS pertencem à classe de modelos guiados por observações, isto é, modelos em que os parâmetros variam no tempo de acordo com um mecanismo que envolve os valores passados do parâmetro e o passado das observações via função score. Os modelos baseados em score fornecem um método de previsão atraente e têm sido empregados com sucesso em uma variedade de aplicações empíricas para previsão de variáveis econômicas e financeiras. Os primeiros resultados comparativos de previsão pontual, estabilidade e densidade preditiva mostram que os modelos GAS, principalmente com distribuição *t-Student*, se mostram superiores em relação aos modelos competidores, porém com maiores erros de previsão médios que as projeções do Boletim Focus, fato esse corroborado por testes de significância estatística. Por conta disso, procurou-se estratégias empíricas para aprimorar a acurácia dos modelos de previsão. Foram adotadas três estratégias: (i) incorporar informações de alta frequência às previsões estimadas pelos modelos individuais via modelo de regressão MIDAS (Mixed Data Sampling) de Ghysels et al. (2004); (ii) a incorporação de expectativas inflacionária via modelo de Black e Litterman (Black e Litterman, 1990); (iii) Por fim, com base em um conjunto de confiança, foi aplicada uma ampla gama de métodos de combinação aos componentes de previsão. Os resultados mostraram que todas essas três etapas foram importantes para ganho de desempenho de previsão. Alguns resultados encontrados por de Castro e Aiube (2022) ratificam os resultados do presente trabalho. Primeiro, os modelos GAS podem ser competitivos em relação às projeções do Boletim Focus e a outros modelos tradicionais na literatura para previsão da inflação brasileira. Segundo, os modelos GAS com distribuição *t-Student* e heterocedástico produziram estimativas de previsão com melhor desempenho em termos de previsão pontual, estabilidade e densidade preditiva. Terceiro, a combinação de modelos foi capaz de gerar previsões de inflação mais robustas quando comparadas a modelos individuais.

Palavras-chave: Previsão de inflação. Generalized Autoregressive Score. Boletim FOCUS.

ABSTRACT

CASTRO, Carlos H. D. C. de. *Inflation time series forecasting using dynamic score-driven models and predictive accuracy improvement methods: the case of Brazil*, 2023. 93 f. Tese (Doutorado em Ciências Econômicas) - Faculdade de Ciências Econômicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

This thesis analyzes monthly forecasts of Brazilian inflation measured by the IPCA. These forecasts are obtained using models based on the GAS (Generalized Autoregressive Score) approach, which was developed by Creal et al. (2008) and Harvey and Chakravarty (2008). We compare these GAS-based forecasts with those published in the Boletim Focus, along with other competing models commonly used in the literature. GAS models fall into the category of observation-driven models, where the parameters vary over time based on a mechanism involving past parameter values and past observations via the score function. Score-based models are an attractive method for forecasting and have been successfully applied in various empirical studies related to economic and financial variables. The initial comparative results regarding point predictions, stability, and predictive density indicate that GAS models, particularly those using Student's t-distribution, outperform competing models. However, it's worth noting that GAS models tend to have slightly higher mean prediction errors than the Focus Bulletin projections, a finding supported by statistical significance tests. To address this, we employ several empirical strategies to enhance forecast accuracy. Three strategies were adopted: (i) incorporating high frequency information to forecasts estimated by individual models via the MIDAS (Mixed Data Sampling) regression model by Ghysels et al. (2004); (ii) the incorporation of inflationary expectations via the Black and Litterman model (Black and Litterman, 1990); (iii) Finally, based on a confidence set, a wide range of combination methods were applied to the forecast components. The results showed that all these three steps were important to gain forecasting performance. Some results found by de Castro and Aiube (2022) ratify the results of the present work. First, the GAS models can be competitive in relation to the Boletim Focus projections and other traditional models in the literature for forecasting Brazilian inflation. Second, the GAS models with Student's t-distribution and heteroskedastic produced forecast estimates with better performance in terms of point forecast, stability and predictive density. Third, the combination of models was able to gene

Keywords: Inflation forecast. Generalized Autoregressive Score. FOCUS Bulletin.

LISTA DE ILUSTRAÇÕES

Figura 1 - Passo a passo da estratégia empírica.....	15
Figura 2 - Evolução dos principais índices de preço do Brasil	21
Figura 3 - IPCA mensal (%) de janeiro de 1996 a dezembro de 2019	21
Figura 4 - Exemplo de uma árvore de regressão	31
Figura 5 - Índice de Preço ao Consumidor Amplo (IPCA) e previsões do Boletim Focus.....	34
Figura 6 - Gráfico da estimativa dos parâmetros via regressão linear de rede elástica.....	38
Figura 7 - Curva de Validação Cruzada	39
Figura 8 - Mapa de calor dos p-valores do teste GW	44
Figura 9 - Erros quadrados cumulativos: IPCA.....	45
Figura 10 - Teste de Racionalidade da Previsão (Rossi e Seknposyan, 2016).....	47
Figura 11 - Teste de flutuação de Giacomini e Rossi (2010)	52
Figura 12 - Desvio padrão da série do IPCA.....	54
Figura 13 - Histograma de PITs para previsão um passo à frente.....	55
Figura 14 - Funções de ponderação baseadas na exponencial Almon lag e Beta lag.....	62

LISTA DE TABELAS

Tabela 1 - Descrição dos principais índices de preços do Brasil.....	19
Tabela 2 - Banco de Dados.....	33
Tabela 3 - Estimativas dos parâmetros do modelo ajustado via Regressão linear de rede elástica para λ_{min} e λ_{1se}	40
Tabela 4 - Abreviações usadas no estudo para fins de representação	41
Tabela 5 - Erros quadráticos médios (EQM) e erros absolutos médios (EAM) de previsão ...	42
Tabela 6 - Avaliação de previsão de densidade fora da amostra de todos os modelos e todos os horizontes de previsão	56
Tabela 7 - AIC, BIC e hAh p-value dos modelos de previsão para diferentes funções de restrição	65
Tabela 8 - Erros de previsão	66
Tabela 9 - Desempenho dos modelos fora da amostra para $t + 1, t + 3, t + 6$ e $t + 12$	69
Tabela 10 - Conjunto de confiança.....	73
Tabela 11 - Breve descrição das principais funções de combinação de previsão	80
Tabela 12 - Previsão combinada de erros absolutos médios e erros quadráticos médios.....	81
Tabela 13 - P-valor do teste Giacomini e White em comparação com o Boletim Focus	83
Tabela 14 - Log-scores médios (logS) e CRPS médios ($\times 10$) para todos os modelos e todos os horizontes de previsão	84

LISTA DE ABREVIATURAS E SIGLAS

AR	autoregressivo
BCB	Banco Central do Brasil
CMN	Conselho Monetário Nacional
DCS	<i>dynamic conditional score</i>
DJIA	Dow Jones Industrial Average
GAS	<i>generalized autoregressive score</i>
IGP-DI	Índice Geral de Preços - Disponibilidade Interna
IGP-M	Índice Geral de Preços - Mercado
INCC	Índice Nacional da Construção Civil
INPC	Índice Nacional de Preços ao Consumidor
IPCA	Índice Nacional de Preços ao Consumidor Amplo
IPC-Br	Índice Preços ao Consumidor
LASSO	<i>Least Absolute Shrinkage and Selection Operator</i>
ML-RF	<i>Machine Learning-Random Forest</i>
RW	Random Walk
TVP-VAR	Vetor Autorregressivo com parâmetros variantes no tempo

SUMÁRIO

INTRODUÇÃO.....	13
1. ENSAIO I – PREVISÃO DA INFLAÇÃO BRASILEIRA USANDO MODELOS ORIENTADOS POR SCORE DINÂMICOS E MODELOS COMPETIDORES.....	16
1.1. Índices de preços brasileiros e o IPCA como referência de inflação.....	19
1.2. Evolução da inflação (IPCA) brasileira.....	21
1.3. Boletim FOCUS.....	23
1.4. Métodos empíricos de previsão.....	24
1.4.1. <u>O modelo GAS (Creal et al. (2013))</u>	24
1.4.1.1. Especificação do modelo.....	25
1.4.1.2. Reparametrização.....	26
1.4.1.3. Estimção por máxima verossimilhança.....	27
1.4.1.4. O modelo GAS com distribuição condicional <i>t-student</i>	27
1.4.2. <u>Modelos Competidores</u>	28
1.4.2.1. Modelos de referência para previsão de inflação.....	29
1.4.2.2. Modelos de previsão baseados em Machine Learning.....	31
1.5. Dados	33
1.6. Seleção de variáveis	35
1.6.1. <u>Regressão linear de rede elástica e LASSO</u>	36
1.6.2. <u>Escolhendo lambda</u>	37
1.7. Avaliação da previsão	41
1.7.1. <u>Previsão pontual</u>	42
1.7.2. <u>Avaliação de previsão na presença de instabilidades</u>	45
1.7.2.1. Teste de racionalidade de flutuação de Rossi e Sekhposyan (2016).....	46
1.7.2.2. Teste de flutuação de Giacomini e Rossi (2010).....	50
1.7.3. <u>Densidade preditiva</u>	54
1.8. Discussão	57
2. ENSAIO II – ESTRATÉGIAS EMPÍRICAS PARA GANHO DE ACURÁCIA DE PREVISÃO.....	59
2.1. Previsão da inflação a partir de dados com frequências mistas	60

2.1.1.	<u>Regressão MIDAS</u>	60
2.1.2.	<u>Restrições funcionais de ponderação</u>	61
2.1.3.	<u>Seleção de especificações e testes de adequação</u>	62
2.1.4.	<u>Previsão</u>	63
2.1.5.	<u>Resultados</u>	64
2.2.	Incorporação de expectativas via modelo Black e Litterman	66
2.2.1.	<u>Apresentação do modelo</u>	67
2.2.2.	<u>Dados</u>	68
2.2.3.	<u>Resultados</u>	69
2.3.	Combinação de modelos de previsão	70
2.3.1.	<u>Conjunto de confiança (MCS)</u>	71
2.3.2.	<u>Previsões combinadas</u>	73
2.3.2.1.	Métodos de Combinação.....	74
2.3.2.2.	Resultados.....	80
	CONCLUSÃO	85
	ANEXO - Teste de raiz unitária: Jan/1996 – Dez/2014 (dentro da amostra).....	93

INTRODUÇÃO

Esta tese analisa as previsões mensais da inflação brasileira, medida pelo IPCA, obtidas por modelos baseados na abordagem GAS (*Generalized Autoregressive Score*) desenvolvido por Creal et al. (2008) e Harvey e Chakravarty (2008) com as previsões divulgadas pelo Boletim Focus. A pesquisa foi dividida em dois ensaios. No ensaio I, foram estimadas as previsões da inflação brasileira usando modelos orientados por score dinâmicos e modelos competidores. Já no Ensaio II, foram aplicadas estratégias empíricas para ganho de acurácia de previsão.

Os modelos GAS pertencem à classe de modelos guiados por observações, isto é, modelos em que os parâmetros variam no tempo de acordo com um mecanismo que envolve os valores passados do parâmetro e o passado das observações via função score. Os modelos baseados em score fornecem um método de previsão atraente e têm sido empregados com sucesso em uma variedade de aplicações empíricas para previsão de variáveis econômicas e financeiras.

Diante de um ambiente rico em dados disponíveis, o número de variáveis preditoras relevantes é potencialmente grande. A abordagem GAS permite que sejam utilizadas variáveis preditoras adicionais na composição do modelo. Assim, no Ensaio I foi aplicado o procedimento de rede elástica e LASSO (*Least Absolute Shrinkage and Selection Operator*), inicialmente desenvolvido por Tibshirani (1996), para selecionar o conjunto de variáveis que pudessem aumentar o poder de precisão do modelo.

Os resultados comparativos do Ensaio I de previsão pontual, estabilidade e densidade preditiva mostram que os modelos GAS, principalmente com distribuição *t-Student*, se mostram superiores em relação aos modelos competidores, porém com maiores erros de previsão médios que as projeções do Boletim Focus, fato esse corroborado por testes de significância estatística. Por conta disso, procurou-se estratégias empíricas para aprimorar a acurácia dos modelos de previsão. Essas estratégias foram aplicadas no Ensaio II.

O primeiro passo do Ensaio II foi aplicar o modelo de regressão MIDAS (*Mixed Data Sampling*) de Ghysels et al. (2004) para incorporar informações de alta frequência às previsões estimadas pelos modelos individuais. A introdução de informações semanais do índice de preço ao consumidor semanal (IPC-S), da variação da cotação Tesouro IPCA+ com Juros Semestrais (NTB051535)¹ e variação internacional do preço do petróleo aumentaram a acurácia dos modelos para horizontes de previsão mais curtos um e três meses à frente, porém tiveram pouca relevância para horizontes mais longos.

A próxima etapa foi a incorporação de expectativas inflacionária de especialistas acadêmicos e do mercado via modelo de Black e Litterman (Black e Litterman, 1990) às novas previsões estimadas pelo modelo de regressão MIDAS. O banco de dados utilizado foi o *surveys of professional forecasters* (SPF) do Banco Central do Brasil. Este banco de dados é formado pelo conjunto de previsões de especialistas que formam o sistema de expectativas do mercado do Banco Central do Brasil. Estes são os dados brutos utilizados pelo Boletim Focus para calcular as projeções divulgadas. Novamente houve ganho de performance nas previsões.

Por fim, com base em um conjunto de confiança, foi aplicada uma ampla gama de métodos de combinação aos componentes de previsão. Após a aplicação de testes de acurácia da previsão pontual e densidades preditiva foi verificado que é possível melhorar a precisão

¹ Notas do Tesouro Nacional série B (NTN-B), papéis do governo federal que pagam a correção da inflação do IPCA mais juros. Fonte: Bloomberg

combinando modelos. Segundo Garcia et al. (2017), a combinação de modelos de previsão de inflação com base em conjuntos de confiança pode alcançar desempenhos preditivos superiores aos modelos individuais.

Os resultados mostraram que os métodos de combinação mais simples produziram os erros médios mais baixos para todos os horizontes de previsão, inclusive sendo superiores ao Boletim Focus para todos os horizontes de previsão.

Alguns resultados encontrados por de Castro e Aiube (2022) ratificam os resultados do presente trabalho. Primeiro, os modelos GAS podem ser competitivos em relação às projeções do Boletim Focus e a outros modelos tradicionais na literatura para previsão da inflação brasileira. Segundo, os modelos GAS com distribuição *t-Student* e heterocedástico produziram estimativas de previsão com melhor desempenho em termos de previsão pontual, estabilidade e densidade preditiva. Terceiro, a combinação de modelos foi capaz de gerar previsões de inflação mais robustas quando comparadas a modelos individuais.

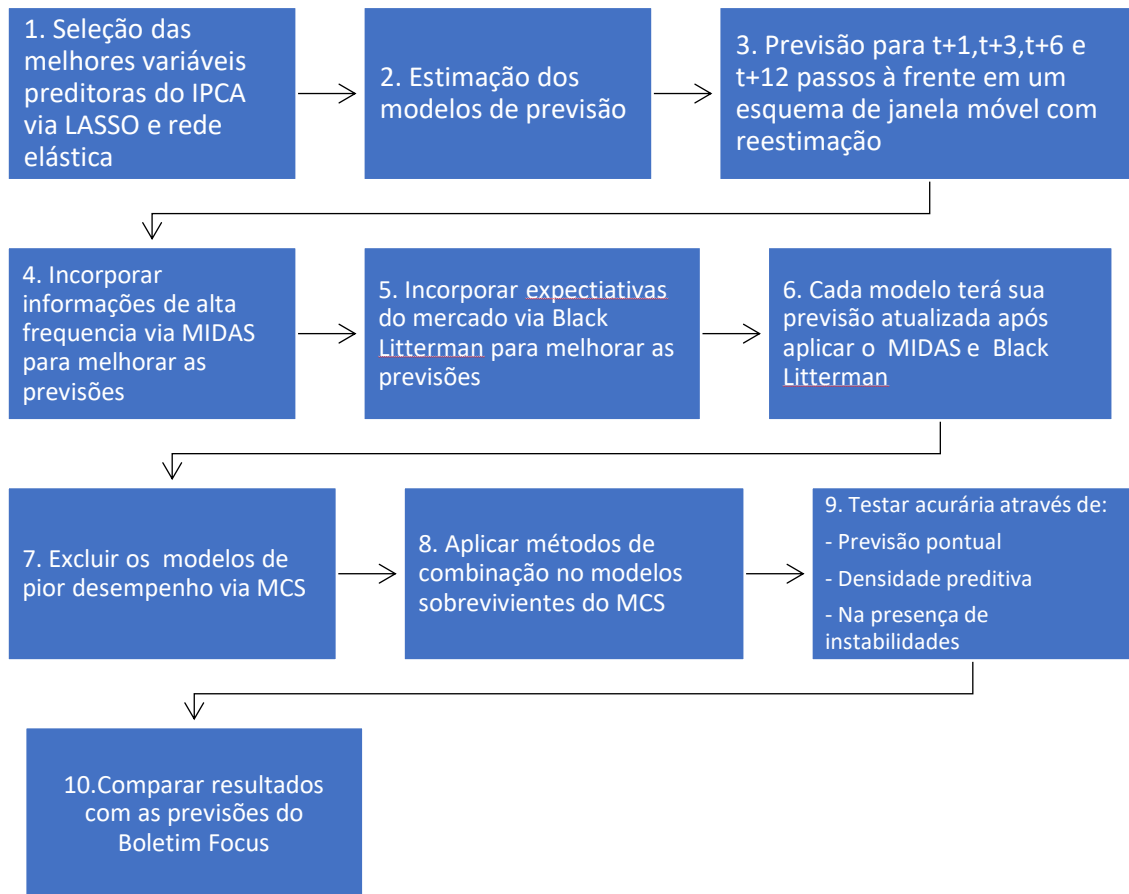
Esta pesquisa contribuiu com a literatura de previsão da inflação brasileira em diversas direções: (i) Diante de uma grande quantidade de dados disponíveis atualmente, foi aplicado um método de seleção de variáveis LASSO e rede elástica para selecionar as variáveis com maior potencial para aumentar a capacidade preditiva dos modelos baseados na abordagem GAS. Os modelos GAS, principalmente com a evolução das inovações com distribuição *t-Student* apresentaram desempenho superior em avaliação de previsão pontual e densidade preditiva em relação aos modelos de referência da literatura. Isso ainda é pouco documentado; (ii) A aplicação da regressão MIDAS utilizando como variáveis predictoras as previsões geradas pelos modelos estimados nesta pesquisa, e outras variáveis de alta frequência, se mostraram uma alternativa útil para o pesquisador que deseja incluir o efeito de variáveis de alta frequência na previsão gerada por modelos que não permitem a estimação com um conjunto de dados de frequências mistas; (iii) A incorporação de informações de expectativas inflacionárias presentes no *surveys of professional forecasters* (SPF) do Banco Central através do modelo de Black e Litterman foi capaz de gerar uma nova inflação esperada com menores erros de previsão médios. Esse experimento em um contexto de previsão de inflação, sendo o IPCA o único ativo, ainda não havia sido documentado na literatura; (iv) Seguir a estratégia de previsão do Boletim Focus utilizando os mesmos dados brutos e aplicar funções de combinação nas previsões se mostraram eficientes para ganho de acurácia preditiva.

Com objetivo central de superar as previsões divulgadas pelo Boletim FOCUS, a estratégia empírica de previsão da inflação brasileira deste estudo se resume nas seguintes etapas:

O primeiro passo do estudo foi a aplicação do método de seleção de variáveis LASSO e rede elástica para selecionar as variáveis com maior potencial para melhorar a capacidade preditiva dos modelos baseados na abordagem GAS. Após selecionadas as melhores variáveis predictoras, foi realizada a estimação de quatro especificações GAS: (i) Normal e homocedástico; (ii) Normal e heterocedástico; (iii) *t-Student* e homocedástico; (iv) *t-Student* e heterocedástico. Além disso, foram estimados um conjunto de modelos competidores mais utilizados na literatura de previsão de inflação brasileira. A previsão de $t+1$, $t+3$, $t+6$ e $t+12$ passos à frente de cada modelo foi realizada em esquema de janela móvel com reestimação. Para melhorar as previsões, foram incorporadas informações de alta frequência às previsões de cada modelo através do modelo MIDAS. Assim, cada previsão gerada por cada modelo foi atualizada, gerando novas previsões. As previsões estimadas pelo modelo MIDAS serão novamente atualizadas pelo modelo Black Litterman através da incorporação de informações de expectativas de mercado. Cada modelo teve sua previsão atualizada duas vezes. Primeiro, pelo modelo MIDAS e segundo, pelo modelo Black Litterman. Após concluídas as atualizações

das previsões, foi utilizado um conjunto de confiança (MCS) para excluir os modelos de pior desempenho. Após definidos os modelos sobreviventes do conjunto de confiança, foram aplicados diversos métodos de combinação nas previsões geradas por esses modelos que não foram excluídos. Cada técnica de combinação gerou uma nova previsão. A acurácia preditiva foi testada por previsão pontual, densidade preditiva e avaliação de desempenho na presença de instabilidades. A última etapa deste estudo foi a comparação dos resultados com as previsões divulgadas pelo Boletim FOCUS. A Figura 1 reporta a estratégia empírica aplicada neste estudo.

Figura 1 - Passo a passo da estratégia empírica



1. ENSAIO I – PREVISÃO DA INFLAÇÃO BRASILEIRA USANDO MODELOS ORIENTADOS POR SCORE DINÂMICOS E MODELOS COMPETIDORES

Previsões acuradas de agregados macroeconômicos são cruciais para tomada de decisão dos agentes (Iversen et al., 2016). Especificamente, a inflação é uma variável econômica difícil de prever. No entanto, previsões confiáveis proporcionam um ambiente econômico mais transparente. Por exemplo, as empresas precisam conhecer a inflação futura para tomar suas decisões de investimento; os investidores usam as previsões de inflação para conhecer seus ganhos reais; os salários são reajustados com base na inflação futura; e contratos como aluguel e hipotecas o utilizam como índice para corrigir preços futuros. Além disso, os Bancos Centrais geralmente precisam de previsões de inflação confiáveis para conduzir a política monetária (metas de inflação).

Mercados emergentes costumam ter inflação mais alta e volátil quando comparados a países desenvolvidos. Para países como o Brasil, que tiveram uma longa e recente experiência de hiperinflação, é ainda mais importante entender os componentes e o comportamento futuro da inflação.

De meados da década de 1980 até 1994, o Brasil sofreu de hiperinflação intercalada com breves períodos de estabilidade devido a programas de reforma monetária malsucedidos. Mesmo depois da estabilização de preços advinda com a implementação do plano Real em 1994 e a adoção de metas de inflação em 1999 para tornar a economia mais atraente e estável, o Brasil passou por diversos períodos de alta inflacionária recente. O regime de metas inflacionárias é uma estratégia de condução da política monetária baseada no anúncio de uma meta para a inflação no início de determinado período que é estabelecida pelo governo e deve ser perseguida pelo Banco Central. Segundo Carrara e Correa (2012), este regime visa a uma maior transparência na condução da política monetária baseada na busca pelo aprimoramento dos canais de comunicação entre o Banco Central e os agentes econômicos para que possa haver um monitoramento e uma avaliação do desempenho da autoridade monetária. Desta forma, essas metas podem coordenar a formação de expectativas inflacionárias dos agentes e a fixação de preços, atuando, assim, como uma âncora nominal tanto para a inflação atual como para as expectativas futuras. Isso é particularmente mais relevante em períodos de crises inflacionárias, como o pós-COVID-19.

Portanto, a previsão da inflação em mercados emergentes é mais relevante do que nas economias desenvolvidas, tornando o exercício de projeção mais difícil e sensível a muitas variáveis. Nesse sentido, muitos especialistas têm se empenhado em fazer melhores previsões, aplicando diversas técnicas econométricas.

Há um crescente corpo de pesquisadores sobre previsão de inflação no Brasil. Arruda et al. (2011) usaram vários modelos lineares e não lineares e a curva de Phillips. Os autores mostraram que alguns modelos não lineares e o modelo autoregressivo simples (AR) produziram erros de previsão menores do que a curva de Phillips. Figueiredo e Marques (2009) usaram modelos heteroscedásticos de memória longa para mostrar que a inflação brasileira tem dependência de longo prazo tanto da média quanto da variância. No entanto, eles não excluíram a importância do componente de AR de curto prazo. A relevância da inflação passada também foi apontada por Kohlscheen (2012). Mais recentemente, Medeiros et al. (2016) e Garcia et al. (2017) consideraram diferentes modelos de alta dimensão para a previsão da inflação brasileira.

Os autores mostraram que as técnicas baseadas no LASSO (*Least Absolute Shrinkage and Selection Operator*) apresentam boa performance.

As duas últimas décadas foram marcadas pelo crescente interesse em modelos com parâmetros variantes no tempo (*Time-varying parameter* - TVP) para a análise de variáveis macroeconômicas, para os quais a literatura propôs uma infinidade de especificações possíveis. Stock e Watson (1996) renovaram o interesse nesta área documentando os ganhos substanciais de previsão dos modelos TVP. Cogley e Sargent (2005), Primiceri (2005) e Stock e Watson (2007) destacaram a importância de permitir a variação temporal na volatilidade, bem como nos coeficientes. Creal et al. (2008) e Harvey e Chakravarty (2008) observaram que muitos dos modelos propostos são difíceis de estimar (em particular a classe de modelos de volatilidade estocástica revisados em Shephard, 2005) e/ou não levam em conta adequadamente a forma da distribuição condicional dos dados. A maioria dos estudos considera modelos TVP sob a suposição de que os erros são normalmente distribuídos. Embora essa suposição seja conveniente, limita a capacidade do modelo de capturar o comportamento da cauda da distribuição, importante característica de séries de tempo de variáveis macroeconômicas.

Creal et al. (2008) e Harvey e Chakravarty (2008) propuseram, portanto, que os parâmetros variantes no tempo sejam conduzidos pelo score da função de densidade condicional. O modelo resultante ficou conhecido como: *score-driven* model, *dynamic conditional score* (DCS) model, ou *generalized autoregressive score* (GAS) model.

Os modelos GAS pertencem à classe de modelos guiados por observações, isto é, modelos em que os parâmetros variam no tempo de acordo com um mecanismo que envolve os valores passados do parâmetro e o passado das observações via função score. Os modelos baseados em score fornecem um método de previsão atraente e têm sido empregados com sucesso em uma variedade de aplicações empíricas para previsão de variáveis econômicas e financeiras. Os desempenhos de previsão dos modelos GAS são investigados em detalhes por Koopman et al. (2016). A abordagem GAS é flexível em termos de especificação da distribuição condicional do termo de inovação, mantendo uma implementação simples. A estimação de parâmetros desconhecidos em modelos GAS pode ser baseada em inferência da verossimilhança analítica, que não requer métodos computacionalmente intensivos ou baseados em simulação.

É possível especificar modelos GAS em formas de componentes, decompondo a média condicional da série em componentes não observáveis, tais como tendência, sazonalidade, ciclo e erro. Utilizando essa metodologia é possível considerar que a fonte de erro da variável de interesse siga outra distribuição além da normal. Isso é interessante para a série temporal utilizada nesse estudo, uma vez que algumas observações podem ser consideradas outliers se tratadas nesse contexto quando na verdade podem ser provenientes de uma distribuição de cauda mais pesada, como a *t-Student*. A metodologia fornece resultados relativos à distribuição assintótica dos estimadores e outras propriedades estatísticas de uma maneira compreensiva e direta, mesmo ao aumentar a complexidade do modelo ao acrescentar restrições aos parâmetros e às variáveis de interesse.

Há uma bibliografia vasta do uso dos modelos GAS para previsão em finanças como Blazsek et al. (2018) para projeção do Dow Jones Industrial Average (DJIA) index e Blasques et al. (2017), Nystrup et al. (2017) e Blazsek et al. (2016) para o S&P 500 index, porém a aplicação em variáveis macroeconômicas ainda é pouco documentada. Alguns exemplos são Delle Monache e Petrella (2017), Gorgi et al. (2019) e de Castro e Aiube (2022) para previsão de inflação e Blasques et al. (2019), Bisaglia e Grigoletto (2020) e Lucas e Zhang (2016) para previsão de taxas de câmbio e dos Santos et al. (2019) para previsão de curvas de juros.

Assim, existem pelo menos três razões pelas quais o modelo GAS é particularmente adequado para previsões de inflação. Em primeiro lugar, modelos autorregressivos univariados simples mostraram funcionar bem no contexto da previsão da inflação (ver Faust e Wright, 2013). Segundo, Pettenuzzo e Timmermann (2015) mostram que os modelos com parâmetros variantes no tempo superam os modelos de parâmetros constantes e que os modelos com mudanças pequenas / frequentes, como é o caso dos modelos utilizados nesta pesquisa, produzem previsões mais precisas do que os modelos cujos parâmetros exibem mudanças grandes / raras. Terceiro, a distribuição dos erros é flexível. Este último ponto é particularmente relevante, pois é bem conhecido, pelo menos desde o trabalho seminal de Engle (1982), que a distribuição da inflação apresenta características não gaussianas.

Este trabalho faz uso da modelagem econométrica baseado no framework do modelo GAS para estimar as projeções do índice de preços brasileiro (IPCA – Índice Nacional de Preços ao Consumidor Amplo). Essa não é apenas a medida de inflação mais usada no Brasil, mas também o índice usado para definir a meta de inflação para a política do Banco Central do Brasil.

A abordagem GAS permite que sejam utilizadas variáveis preditoras na composição do modelo. Assim, foi aplicado o procedimento de rede elástica e LASSO para selecionar um conjunto de variáveis que pudessem aumentar o poder de predição do modelo.

Além da aplicação do modelo GAS para prever a inflação brasileira, foram estimados modelos competidores da literatura de previsão para realizar a comparação de performance. Dentre estes, estão os modelos: Random Walk (RW), Auto-regressivo (AR), família GARCH (Engle (1982) e Bollerslev (1986)), Curva de Phillips, VAR (Vetor Auto-Regressivo), VAR com parâmetros variantes no tempo, dois modelos baseados em *machine learning*: *ML-Random Forest* (Breiman, 2001) e *ML-LASSO* (Tibshirani, 1996). Para verificar a robustez e precisão de todos os modelos de previsão, foram avaliadas as previsões pontuais e de densidade.

Os resultados também foram comparados a compilação de projeções de especialistas catalogadas pelo Banco Central do Brasil (BCB). Essas projeções são obtidas do Boletim Focus, que contém previsões para diversas variáveis da economia brasileira. O Boletim Focus é baseado em uma pesquisa online de projeções de mais de uma centena de analistas profissionais (associados a bancos, gestores de fundos e empresas de consultoria, etc.). O relatório é elaborado para subsidiar o regime de metas de inflação e é divulgado semanalmente pelo BCB. Será utilizada a mediana dessas expectativas de mercado sobre a inflação brasileira, conforme divulgado no Boletim Focus. Além disso, o Boletim Focus também publica as expectativas do “Top 5”, as previsões dos cinco agentes mais precisas em períodos anteriores. Nesta pesquisa será chamado de “Top5Focus”. Além da inflação, o Boletim divulga expectativas sobre PIB, produção industrial, câmbio e outras variáveis. Todas essas informações são utilizadas pelo BCB para avaliar sua política monetária.

Foram estimadas previsões a partir dos modelos propostos para horizontes de um, três, seis e doze meses à frente. Os erros das previsões do Boletim Focus dominaram todos os modelos individuais. Em termos de erros quadráticos médios e erro absoluto médio, entretanto, sua performance tornou-se inferior à medida que o horizonte de previsão aumentava.

O Ensaio I está organizado da seguinte forma: a Seção 2 descreve os principais índices de preço utilizados no Brasil; a Seção 3 apresenta um breve histórico do Índice de Preços ao Consumidor (IPCA) brasileiro; a Seção 4 explica como é construído o Boletim Focus, a Seção 5 apresenta o modelo GAS e as estratégias empíricas competidoras; a Seção 6 descreve o conjunto de dados; Seção 7 apresenta a estratégia para selecionar as variáveis preditores do modelo GAS; a Seção 8 apresenta e discute os resultados.

1.1. Índices de preços brasileiros e o IPCA como referência de inflação

Inflação significa um aumento generalizado dos preços na economia. Para medi-la são construídos índices de preços, que tomam uma média de diversos preços de modo a resumi-los em um único número. Os índices de preços podem diferir de várias maneiras, destacando-se as diferenças na cesta de bens e serviços tomada como referência. Por exemplo, pode-se construir índices de preços ao consumidor, índices de preços ao produtor, índices de custos de produção, etc, a depender do objetivo. Mesmo índices de preços ao consumidor podem diferir entre si, pois cada família tem sua própria cesta de consumo, e um índice pode ser desenhado para refletir o custo de vida para um ou outro grupo de famílias. Pode haver, por isso, percepções distintas entre o que o cidadão vê no seu orçamento e o que aparece em um dado índice.

O Índice Nacional de Preços ao Consumidor Amplo (IPCA) é calculado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), com coleta, em geral, do dia 1 a 30 do mês de referência. Pela Resolução 2.615, de 30 de junho de 1999, ficou determinado pelo CMN (Conselho Monetário Nacional) que o índice a ser utilizado como referência para o regime de metas de inflação seria o IPCA, calculado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) desde dezembro de 1979. Mede o preço de uma cesta de consumo representativa para famílias com renda de 1 a 40 salários-mínimos, em 13 áreas geográficas: regiões metropolitanas de Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Vitória, Rio de Janeiro, São Paulo, Curitiba, Porto Alegre, além do Distrito Federal e dos municípios de Goiânia e Campo Grande.

O processo inflacionário entre os anos 70 e meados de 90 reforçou a necessidade de se contar com maior variedade de índices. Por um lado, a inflação alta e volátil fez com que a evolução dos diferentes preços diferisse ainda mais entre si, levando à necessidade de índices de preços mais específicos para cada propósito. Por exemplo, o INPC foi criado para refletir o custo de vida de trabalhadores urbanos e passou a ser utilizado como parâmetro de reajuste em dissídios salariais.

Por outro, a inflação alta e volátil também tornou necessário um acompanhamento mais frequente da evolução dos preços. Isso se refletiu na criação do IGP-M, com período de coleta de preços distinto ao do já existente IGP-DI, o que permitiu ao mercado contar com um índice divulgado no último dia do mês para a correção de contratos referentes a operações financeiras e correções de balanços. A descrição dos índices de preços está apresentada na Tabela 1.

Tabela 1 - Descrição dos principais índices de preços do Brasil (continua)

Índices de Preços		Descrição	Fonte
IPCA	Índice Nacional de Preços ao Consumidor Amplo	É o Índice Nacional de Preços ao Consumido Amplo que reflete o custo de vida médio das 11 maiores cidades com renda mensal mínima de 40 salários-mínimos. O IPCA é utilizado pelo Banco Central do Brasil para decisão de política monetária.	IBGE
IPCA-15	Índice Nacional de Preços ao Consumidor Amplo -15	Difere do IPCA, apenas no período de coleta que abrange, em geral, do dia 16 do mês anterior ao 15 do mês de referência e na abrangência geográfica.	IBGE

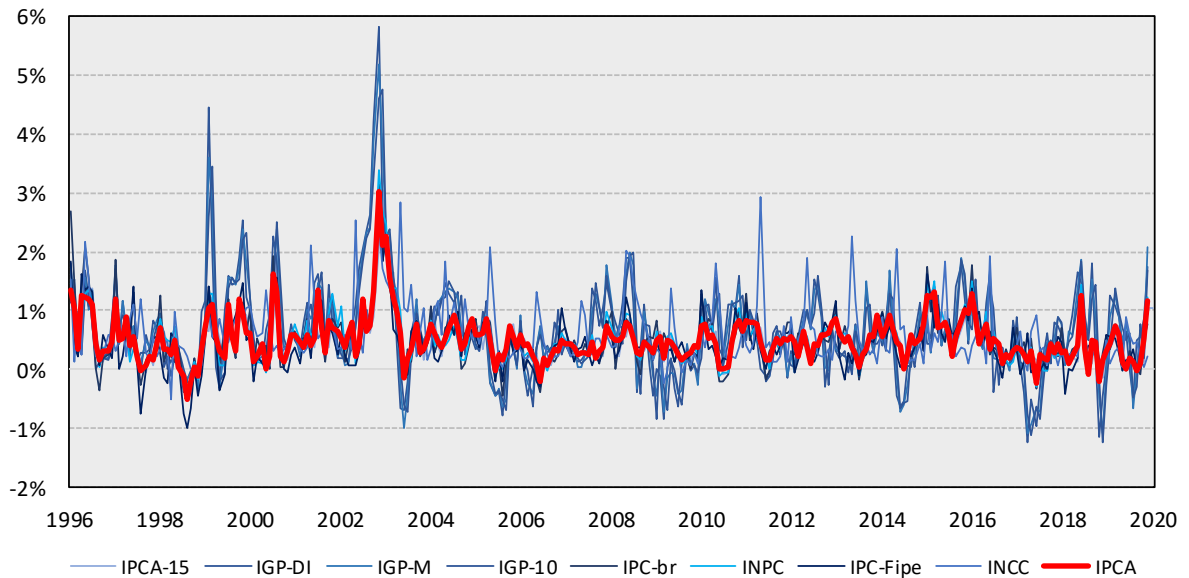
Tabela 1 - Descrição dos principais índices de preços do Brasil (conclusão)

IGP-DI	Índice Geral de Preços - Disponibilidade Interna	Preços mensais de todo o processo produtivo: matérias-primas agrícolas e industriais, produtos intermediários e bens e serviços finais e preços de construção. É parte da cesta que corrige os preços de telefonia.	FGV
IGP-M	Índice Geral de Preços - Mercado	Preços do comércio no atacado, no varejo e na construção civil, pesquisados entre o dia 21 do mês anterior e 20 do mês de referência. É usado na correção de contratos de aluguel e tarifas de serviços públicos.	FGV
IGP-10	Índice Geral de Preços - 10	É um indicador do movimento de preços que cobre todo o processo produtivo, desde preços de matérias-primas agrícolas e industriais, passando pelos preços de produtos intermediários até os de bens e serviços finais. O IGP-10 mede a evolução de preços no período compreendido entre o dia 11 do mês anterior e o dia 10 do mês de referência.	FGV
IPC-Br	Índice Preços ao Consumidor	Mede mensalmente a variação de preços de bens e serviços que compõem as despesas habituais das famílias com renda entre 1 e 33 salários-mínimos mensais.	FGV
IPC-S	Índice de Preços ao Consumidor Semanal	Mede quadrissemanalmente a variação de preços de bens e serviços que compõem as despesas habituais das famílias com renda entre 1 e 33 salários-mínimos mensais.	FGV
IPC-Fipe	Índice de preço ao consumidor - Fundação Instituto de Pesquisas Econômicas	Calcula semanalmente os preços de 468 itens consumidos por famílias de que recebem entre 0 e 10 salários na cidade de São Paulo.	FGV
INPC	Índice Nacional de Preços ao Consumidor	Variação do custo médio das famílias com rendimento familiar médio entre 1 e 5 salários-mínimos. Indica as variações de preços nos grupos mais sensíveis, que gastam todo rendimento em consumo corrente (alimentação, remédio, etc.).	IBGE
INCC	Índice Nacional da Construção Civil	Indicador que apura a evolução dos custos no setor da construção habitacional, um dos termômetros do nível de atividade da economia.	FGV

Fonte: IBGE: Instituto Brasileiro de Geografia e Estatística, FGV: Fundação Getúlio Vargas, FIPE: Fundação Instituto de Pesquisas Econômicas.

A Figura 2 apresenta a trajetória comparativa dos índices de preços do Brasil. O índice de referência, IPCA, está colorido na vermelha, já os outros índices estão na cor em tons de azul. Observe que o IPCA é um dos índices de inflação menos voláteis. Isso se deve pela composição dos produtos da cesta de bens e serviços e características das famílias que o índice abrange. Comparativamente ao IPCA, o IGP-M é visto como mais volátil dado que tem um peso maior dos preços das matérias-primas e dos insumos, que incidem sobre a inflação no atacado, e assim mais sensíveis às variações dos preços das commodities e câmbio.

Figura 2 - Evolução dos principais índices de preço do Brasil

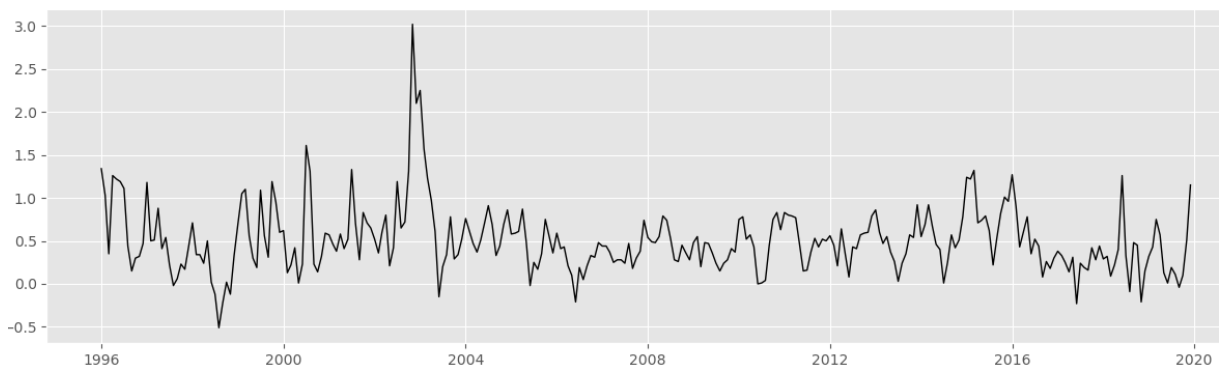


Fonte: O Autor, 2022. Elaboração própria com dados do IBGE e FGV.

1.2. Evolução da inflação (IPCA) brasileira

O índice de preços ao consumidor amplo (IPCA) brasileiro passou por diferentes fases desde sua criação em 1991. O período de 1994-96 corresponde à implantação do Plano Real, quando a âncora cambial possibilitou o fim do longo período de inflação elevada no Brasil e contribuiu para a aceleração do PIB, que teve crescimento médio de 4.0% a.a. (ao ano). Nesse período a política fiscal foi relativamente expansionista, houve aumento do déficit em conta corrente e dependência crescente da entrada de capitais externos para manter a âncora cambial. A Figura 3 apresenta a evolução do IPCA mensal brasileiro.

Figura 3 - IPCA mensal (%) de janeiro de 1996 a dezembro de 2019



Fonte: O Autor, 2022. Elaboração própria com base nos dados do IBGE.

O período 1997-99 foi marcado pelo enfraquecimento da indústria nacional, além dos efeitos das crises internacionais sobre o balanço de pagamentos. A pressão inflacionária

causada pelos efeitos da crise cambial de 1999 foi minimizada por uma política monetária restritiva (taxa básica de juros extremamente alta – Selic, Sistema Eletrônico de Liquidação e Custódia). Isso resultou em uma desaceleração do crescimento do PIB e um aumento na taxa de desemprego.

Os anos 2000-02 foram marcados pela implantação do tripé macroeconômico, com a consolidação do regime de metas de inflação, câmbio flutuante e metas de resultado primário. Aliado a isso e à política monetária expansionista, o crescimento do PIB acelerou para 3,0% a.a. e a inflação aumentou significativamente para 8,7% a.a. no triênio. O choque inflacionário decorreu da combinação de restrições de oferta (o racionamento de energia elétrica de 2001) com a fragilidade externa da economia (reservas internacionais reduzidas e maxi-depreciação cambial em 2002). Vale um adendo para o segundo semestre de 2002, caracterizado pela alta do dólar e choque inflacionário, resultado da expectativa negativa do mercado frente ao favoritismo do candidato à presidência Lula, isto ficou conhecido como “efeito Lula”. Observe que houve um pico de 3.02% a.m. (ao mês).

Entre 2003-05, o governo adotou medidas fiscais e monetárias muito restritivas para combater o desequilíbrio externo e a inflação crescente dos anos anteriores (Barbosa-Filho, 2008). Então, em 2006-08, o ambiente externo extremamente favorável permitiu ao governo acumular mais reservas internacionais e reduzir a taxa de juros real devido ao aumento dos preços das commodities. O governo adotou uma política fiscal expansionista com orçamento equilibrado. Esse período foi um dos mais positivos para a economia brasileira, resultando em crescimento constante do PIB, redução da taxa básica de juros, menor desemprego, câmbio forte e controle da inflação.

O quinto período corresponde à “inflexão da política econômica do governo Lula”, em 2006-08 (Barbosa-Filho, 2015). Nessa fase também houve um ambiente externo extremamente favorável, que contribuiu para que o governo acumulasse mais reservas internacionais e reduzisse a taxa real de juros devido ao efeito do aumento dos preços de commodities. O governo adotou uma política fiscal expansionista com orçamento equilibrado, direcionando o aumento de arrecadação principalmente para transferência de renda e investimentos. O PIB de fato acelerou, para o ritmo mais alto dos últimos vinte anos, enquanto a inflação caiu para o seu nível também mais baixo das últimas décadas.

O desempenho nos anos seguintes, 2009-11, sofreu os efeitos da crise financeira internacional de 2008 e da política anticíclica adotada pelo governo brasileiro (Barbosa-Filho, 2015). Com a queda dos preços das commodities e a desaceleração do comércio mundial, os anos seguintes foram marcados pelo ajuste macroeconômico inevitável com a implementação de medidas fiscais e monetárias contracionistas. Passados os efeitos mais agudos da crise, em 2011, o governo iniciou uma consolidação fiscal e monetária, revertendo parte dos estímulos dos anos anteriores.

O sétimo período corresponde à deterioração da política econômica em 2012-14 (Dweck e Teixeira, 2017). Do lado externo, o cenário mudou a partir de 2012, com queda dos preços de commodities e a desaceleração do comércio mundial. Do lado interno, houve esgotamento das medidas de estímulo adotadas no período anterior, bem como manutenção de projetos públicos superdimensionados e intervenção excessiva em alguns mercados regulados. Essa combinação de fatores aumentou a incerteza econômica.

O oitavo período, em 2015-17, corresponde ao inevitável ajuste macroeconômico e à crise política e institucional. Esta fase começou com medidas fiscais e monetárias contracionistas, em 2015, baseadas no aumento do resultado primário estrutural, elevação da taxa de juros real e liberalização dos preços monitorados. O impacto inicial negativo dessas medidas foi amplificado por dois choques: uma nova queda abrupta dos preços das commodities,

no âmbito externo, e os efeitos iniciais da Operação Lava Jato² sobre o nível de atividade e emprego, no âmbito interno. O acirramento da polarização política impediu o ajuste da política econômica a esses dois choques recessivos até o início de 2016, quando a mudança de governo finalmente permitiu a flexibilização da política fiscal. Os choques adversos também se dissiparam a partir do final de 2016, promovendo a estabilização da renda e a redução da inflação em 2017.

No ano de 2018, entretanto, a inflação acelerou. Pode-se atribuir o repique inflacionário à greve dos caminhoneiros, à desvalorização do real frente ao dólar e ao efeito estimulante da redução da Selic. Além disso, com o aumento da inflação para o centro da meta do Banco Central, e com manutenção constante da Taxa Selic nominal, a taxa real de juros caiu ao longo de 2018, intensificando ainda mais o estímulo monetário na economia. Em 2019, apesar da recuperação da economia, mesmo que ainda relativamente discreta, a inflação subiu devido a choques de oferta.

1.3. Boletim FOCUS

O Boletim Focus resume as estatísticas calculadas das expectativas de mercado. O relatório traz a evolução gráfica e o comportamento das projeções para as principais variáveis macroeconômicas como índices de preços, atividade econômica, câmbio, taxa Selic, entre outros indicadores.

O Boletim Focus é visto como uma importante ferramenta de apoio ao planejamento econômico de cidadãos e empresas, pois com base nas expectativas dos agentes de mercado é possível tomar decisões de curto, médio e longo prazo mais seguras e embasadas em uma premissa externa e consistente.

Para guiar as expectativas dos agentes, é importante que o Banco Central do Brasil (BCB) seja transparente quanto a estratégia de política monetária. O Brasil segue o regime de metas de inflação que âncora as expectativas dos agentes econômicos do mercado, pois estes orientam o processo de formação de preços na economia.

Com esta clareza informacional, o BCB ganha maior credibilidade e com isso, os reajustes de preços da economia tendem a se aproximar da meta estabelecida.

A área responsável do BCB por recolher os dados e por preparar o Boletim Focus é a Gerência Executiva de Relacionamento com Investidores (Gerin), que foi criada em abril de 1999, como parte do arcabouço do regime monetário de metas de inflação. Seu objetivo é aperfeiçoar a comunicação entre o BCB e o setor privado, com foco nos investidores domésticos e externos. A Gerin produz seus próprios relatórios, que inclui:

- (i) Focus-Relatório de Mercado ou Boletim Focus: apresentação dos resultados da pesquisa de expectativas de mercado, um levantamento diário das previsões de cerca de 90 bancos e empresas não-financeiras para a economia brasileira, publicado toda segunda-feira;

² É uma operação da Polícia Federal do Brasil que investiga esquemas bilionários de corrupção envolvendo empresas estatais como a Petrobras, empreiteiras e políticos de diferentes partidos.

- (ii) Séries Temporais das Expectativas de Mercado: planilhas atualizadas semanalmente, contendo os dados diários das expectativas de mercado para as principais variáveis da economia desde 2001. Base de dados utilizada para construir o Boletim Focus;
- (iii) Top5Focus: classificação mensal das instituições com melhores previsões dentre aquelas participantes da pesquisa de mercado;

Através do ranking Top5Focus, o Banco Central do Brasil criou uma forma de incentivar e aprimorar a capacidade de previsão dos agentes econômicos participantes da pesquisa que compõe o Boletim Focus. De acordo com Yoshihiro (2006), este ranking tem como prioridade classificar as cinco principais instituições baseada nos índices de acerto. Ou seja, as medianas das variáveis das expectativas projetadas pelas cinco instituições que mais acertarem, ou que estiverem mais próximas ao valor realizado sobre os índices pesquisados, terão seus nomes publicados no relatório chamado de “Top 5 no Relatório de Mercado”. Essas instituições que se colocam neste ranking ganham mais credibilidades sobre seus relatórios no mercado, atraindo cada vez mais clientes. Com isso ocorre uma nova maneira de precificar o trabalho dos analistas e pesquisadores do mercado. Muitas instituições tomam esta classificação para precificar o trabalho de seus pesquisadores, retribuindo o bom trabalho em forma de bônus e benefícios.

“No ranking de curto prazo, avalia-se a precisão das projeções com defasagem de um mês, nos últimos seis meses. O ranking de médio prazo considera a precisão das projeções dos últimos seis meses, sempre para o mês de referência. O ranking de longo prazo considera a precisão das projeções informadas em doze meses para o indicador anual. As instituições Top Five de curto e médio prazo são anunciadas mensalmente, enquanto as instituições Top Five de longo prazo são anunciadas uma vez por ano, em janeiro.” (Yoshihiro, 2005).

1.4. Métodos empíricos de previsão

Esta Seção descreve os métodos usados neste trabalho para prever a inflação futura. Considera-se uma abordagem de previsão em que os h passos à frente (meses) de inflação, π_{t+h} , são modelados como uma função de um conjunto de preditores medidos no tempo t .

1.4.1. O modelo GAS (Creal et al. (2013))

A família de modelos GAS³ é classificada como pertencente aos modelos orientados por observação, a qual permite aos parâmetros a possibilidade de variarem no tempo, através da evolução de uma distribuição condicional (preditiva). A atualização desses parâmetros considera informações da densidade preditiva via o uso da função score.

³ Nesta pesquisa chamaremos a abordagem GAS de “modelos GAS”, apesar de não ser essencialmente um modelo empírico.

1.4.1.1. Especificação do modelo

Seja $\pi_t \in \mathbb{R}^N$ um vetor N -dimensional no tempo t com distribuição condicional:

$$\pi_t | \pi_{1:t-1} \sim p(\pi_t; \boldsymbol{\theta}_t) \quad (1)$$

Em que $\pi_{1:t-1} \equiv (\pi'_1, \dots, \pi'_{t-1})'$ contém os valores passados de π_t até o tempo $t - 1$ e $\boldsymbol{\theta}_t \in \Theta \subseteq \mathbb{R}^J$ é um vetor de parâmetros variantes no tempo que caracteriza a função densidade $p(\cdot)$ e depende apenas de $\pi_{1:t-1}$ e um conjunto de parâmetros estáticos adicionais ξ , ou seja, $\boldsymbol{\theta}_t \equiv \boldsymbol{\theta}(\pi_{1:t-1}, \xi)$ para todo t . A principal característica dos modelos de GAS é que a evolução no vetor de parâmetros variáveis no tempo $\boldsymbol{\theta}_t$ é dirigido pela pontuação da distribuição condicional definida em (1), juntamente com um componente autorregressivo:

$$\boldsymbol{\theta}_t \equiv \kappa + \sum_{i=1}^p A_i s_{t-i} + \sum_{j=1}^q B_j \boldsymbol{\theta}_{t-j} \quad (2)$$

onde κ é um vetor de constantes, A_i e B_j são matrizes de coeficientes com dimensões coletadas em ξ , sendo $i = 1, \dots, p$ e $j = 1, \dots, q$, enquanto s_t é um vetor proporcional ao score de (1):

$$s_t \equiv \mathbf{S}_t(\boldsymbol{\theta}_t) \nabla_t(\pi_t, \boldsymbol{\theta}_t)$$

A matriz \mathbf{S}_t é uma matriz $J \times J$ escalonada positiva definida conhecida no tempo t e $\nabla_t(\pi_t, \boldsymbol{\theta}_t) \equiv \frac{\partial \log p(\pi_t, \boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t}$ é o score de (1) avaliado em $\boldsymbol{\theta}_t$. Creal et al. (2013) sugerem que a matriz \mathbf{S}_t seja ajustada para uma potência $\gamma > 0$ do inverso da matriz de informação de $\boldsymbol{\theta}_t$ para contabilizar a variância de ∇_t . Mais precisamente:

$$\mathbf{S}_t(\boldsymbol{\theta}_t) \equiv \mathfrak{I}_t(\boldsymbol{\theta}_t)^{-\gamma},$$

Sendo

$$\mathfrak{I}_t(\boldsymbol{\theta}_t) \equiv E_{t-1}[\nabla_t(\pi_t, \boldsymbol{\theta}_t) \nabla_t(\pi_t, \boldsymbol{\theta}_t)'] \quad (3)$$

onde E_{t-1} denota a expectativa com respeito à distribuição condicional de π_t dado $\pi_{1:t-1}$. O parâmetro γ é fixo e geralmente assume um valor no conjunto $\{0, \frac{1}{2}, 1\}$, quando $\gamma = 0$, $\mathbf{S}_t = I$ e não há escalonamento⁴. Se $\gamma = 1$ (ou $\gamma = \frac{1}{2}$), o score condicional $\nabla_t(\pi_t, \boldsymbol{\theta}_t)$ é pré-multiplicado pelo inverso de (ou a raiz quadrada de) sua matriz de covariância $\mathfrak{I}_t(\boldsymbol{\theta}_t)$.

⁴ Denota-se por I a matriz de identidade de tamanho apropriado.

Considerando a suposição paramétrica dos modelos GAS, as previsões são geralmente dadas na forma de densidade preditiva, ou seja, a distribuição de $\pi_{T+h} | \pi_{1:T}$ para $h \geq 1$. Conhecendo a densidade preditiva, é possível extrair quaisquer elementos relevantes, como o valor futuro esperado ou (co-) variância. Para os modelos GAS, a distribuição preditiva de um passo à frente ($h = 1$) está analiticamente disponível, embora precise ser estimada por simulação no multi-passos à frente caso ($h > 1$). Blasques et al. (2017) discutem em detalhes as propriedades das previsões no framework GAS.

1.4.1.2. Reparametrização

Em (2), vetor de parâmetros θ_t tem uma especificação linear e, portanto, é ilimitado. Na prática, o espaço de parâmetros de θ_t é frequentemente restrito ($\Theta \subset \mathbb{R}^J$). Por exemplo, quando se modelam os parâmetros de variância (*scale*) de uma distribuição *t-Student*, é necessário garantir sua positividade. Mesmo que isto pode ser resolvido impondo restrições ao ζ (como é feito no modelo GARCH; veja Bollerslev, 1986), a solução padrão sob a abordagem GAS é usar uma função *link* (geralmente não linear) $\Lambda(\cdot)$ que mapeia $\tilde{\theta}_t \in \mathbb{R}^J$ em θ_t e onde $\tilde{\theta}_t \in \mathbb{R}^J$ tem especificação dinâmica linear de (2). Especificamente, seja $\Lambda : \mathbb{R}^J \rightarrow \Theta$ uma função de mapeamento duplamente diferenciada e valorizada por vetor, de modo que $\Lambda(\tilde{\theta}_t) = \theta_t$. A equação de atualização para θ_t é então dada por:

$$\theta_t \equiv \Lambda(\tilde{\theta}_t) \quad (4)$$

$$\tilde{\theta}_t \equiv \kappa + A\tilde{s}_t + B\tilde{\theta}_{t-1}$$

onde, $\tilde{s}_t \equiv \tilde{S}_t(\tilde{\theta}_t)\nabla_t(\pi_t, \theta_t)$ e $\tilde{V}_t(\pi_t, \theta_t)$ representa o score de (1) em relação a $\tilde{\theta}_t$, e consequentemente, $\tilde{S}_t(\tilde{\theta}_t)$ pode depender da matriz de informação de $\tilde{\theta}_t$ dado por $\tilde{\mathfrak{I}}_t(\tilde{\theta}_t)$. Denota-se a Matriz Jacobiana de $\Lambda(\cdot)$ avaliada em $\tilde{\theta}_t$ como:

$$J(\tilde{\theta}_t) \equiv \frac{\partial \Lambda(\tilde{\theta}_t)}{\partial \tilde{\theta}_t}$$

Então, as seguintes relações se mantêm:

$$\tilde{V}_t(\pi_t, \tilde{\theta}_t) = J(\tilde{\theta}_t)' \nabla_t(\pi_t, \theta_t),$$

$$\tilde{\mathfrak{I}}_t(\tilde{\theta}_t) = J(\tilde{\theta}_t)' \mathfrak{I}_t(\theta_t) J(\tilde{\theta}_t)$$

Desta forma, quase todas as restrições não lineares podem ser facilmente tratadas através da definição de uma função de mapeamento adequado $\Lambda(\cdot)$ e sua matriz Jacobiana $J(\cdot)$ associada. Os coeficientes a serem estimados estão reunidos em $\zeta \equiv (\kappa, A, B)$ e estimados através da maximização numérica da função (log-) verossimilhança.

1.4.1.3. Estimação por máxima verossimilhança

Uma propriedade útil dos modelos de GAS é que, dada a informação passada e o vetor de parâmetros estáticos ξ , o vetor de parâmetros variantes no tempo, θ_t , é perfeitamente previsível e a função de verossimilhança de erro de previsão pode ser facilmente avaliada através da decomposição do erro de previsão. Mais precisamente, para uma amostra de T realizações de π_t , coletado em $\pi_{1:T}$, o vetor de parâmetros ξ pode ser estimado pela máxima verossimilhança (ML) como a solução de:

$$\hat{\xi} \equiv \arg \max_{\xi} \mathcal{L}(\xi; \pi_{1:T}) \quad (5)$$

onde

$$\mathcal{L}(\xi; \pi_{1:T}) \equiv \log p(\pi_1; \theta_1) + \sum_{t=2}^T \log p(\pi_t; \theta_t)$$

com $\theta_1 \equiv (I - B)^{-1}\kappa$, e, para $t > 1$, $\theta_t \equiv \theta(\pi_{1:t-1}; \xi)$. Note a dependência do θ_t em ξ e $\pi_{1:t-1}$.

1.4.2. O modelo GAS com distribuição condicional t -Student

Considerando o caso em que a distribuição da variável $\pi_t \in \mathfrak{R}$, condicionalmente em $\pi_{1:t-1}$, é t -Student com média μ_t , desvio-padrão $\phi_t \in \mathbb{R}^+$, e $\nu_t > 2$ graus de liberdade⁵, ou seja, $\theta_t = (\mu_t, \phi_t^2, \nu_t)'$ e:

$$p(\pi_t; \theta_t) \equiv \frac{\Gamma\left(\frac{\nu_t + 1}{2}\right)}{\Gamma\left(\frac{\nu_t}{2}\right) \phi_t \sqrt{\pi \nu_t}} \left(1 + \frac{(\pi_t - \mu_t)^2}{\nu_t \phi_t^2}\right)^{-\frac{\nu_t + 1}{2}} \quad (6)$$

O score correspondente à distribuição t -Student tem a vantagem de atenuar o efeito de observações extremas, sendo mais adequada para aderir a dados empíricos com caudas pesadas. Isso tem sido utilizado por Creal et al. (2013) e Lucas e Zhang (2016) sob o nome GAS-t, e por Harvey (2013) e Harvey e Luati (2014) sob o nome Beta-t-EGARCH.

Diferenciando o logaritmo de (6) com respeito a θ_t leva ao vetor de score $\nabla_t(\pi_t, \theta_t) = (\nabla_t^\mu, \nabla_t^{\phi^2}, \nabla_t^\nu)'$, com:

$$\nabla_t^\mu \equiv \frac{(\nu_t + 1)(\pi_t - \mu_t)}{\nu_t \phi_t^2 + (\pi_t - \mu_t)^2}$$

⁵ Note que o parâmetro graus de liberdade é assumido como um número real maior que dois, o que torna o cálculo da derivada parcial simples.

$$\nabla_t \phi^2 \equiv \frac{1}{2\phi_t^2} \left[\frac{(v_t + 1)(\pi_t - \mu_t)^2}{v_t \phi_t^2 + (\pi_t - \mu_t)^2} - 1 \right]$$

$$\nabla_t^v \equiv \frac{1}{2} \psi \left(\frac{v_t + 1}{2} \right) - \frac{1}{2} \psi \left(\frac{v_t}{2} \right) - \frac{1}{2v_t} - \frac{1}{2} \log \left(1 + \frac{(\pi_t - \mu_t)^2}{v_t \phi_t^2} \right) + \frac{(v_t + 1)(\pi_t - \mu_t)^2}{2v_t [v_t \phi_t^2 + (\pi_t - \mu_t)^2]}$$

onde $\psi(\cdot)$ é a uma função Digamma⁶. Sem perda de generalidade, considera-se o caso em que $\gamma = 0$ sem reparametrização, ou seja, $\theta_t = \tilde{\theta}_t$. Os resultados, quando $\gamma \neq 0$ e uma função mapeamento $\Lambda(\cdot)$ para θ_t é introduzida, são qualitativamente os mesmos. Claramente, que controles para a resposta a observações extremas no score condicional $\nabla_t(\pi_t, \theta_t)$ é o grau de liberdade parâmetro v_t . Quando v_t é pequeno, digamos $v_t = 3$, a distribuição condicional do π_t tem uma massa de probabilidade relativamente maior na cauda, o que significa que as observações extremas, que são consideradas *outliers* sob a distribuição condicional Gaussiana, são mais propensas a serem observadas.

Se introduzirmos a seguinte função de mapeamento para o vetor irrestrito do parâmetro $\tilde{\theta}_t = (\tilde{\mu}_t, \tilde{\phi}_t, \tilde{v}_t)'$:

$$\Lambda(\tilde{\theta}_t) \equiv \begin{cases} \mu_t \equiv \tilde{\mu}_t \\ \phi_t^2 \equiv \exp(\tilde{\phi}_t^2) \\ v_t \equiv \exp(\tilde{v}_t) + c \end{cases}$$

então a etapa de atualização do GAS para θ_t quando $\gamma = 0$ toma a forma:

$$\theta_{t+1} \equiv \Lambda(\tilde{\theta}_{t+1}) \tag{7}$$

$$\tilde{\theta}_{t+1} \equiv \kappa + A\mathcal{J}(\tilde{\theta}_t)' \nabla_t(\pi_t, \theta_t) + B\tilde{\theta}_t$$

onde $\kappa \equiv (\kappa_\mu, \kappa_\phi, \kappa_v)'$, $A \equiv \text{diag}(\alpha_\mu, \alpha_\phi, \alpha_v)$ e $B \equiv \text{diag}(\beta_\mu, \beta_\phi, \beta_v)$. Neste caso particular, a matriz Jacobiana $\mathcal{J}(\tilde{\theta}_t)$ toma a forma:

$$\mathcal{J}(\tilde{\theta}_t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp(\tilde{\phi}_t^2) & 0 \\ 0 & 0 & \exp(\tilde{v}_t) \end{pmatrix}$$

Restrições na evolução dos parâmetros de GAS podem ser facilmente consideradas ao fixar os valores dos elementos A e B. Por exemplo, se a restrição $v_t = v$ tiver que ser imposta, então será definido $a_v = b_v = 0$ durante a maximização da verossimilhança.

1.4.2. Modelos Competidores

⁶ Derivada logaritma da função gama

1.4.2.1. Modelos de referência para previsão de inflação

Muitos autores têm utilizado modelos de referência para prever a inflação brasileira (Costa Filho, 2016; Arruda et al., 2011; Triches e Feijó, 2017). Foram selecionados os modelos mais utilizados como concorrentes aos modelos GAS.

a) Random Walk (RW)

$$\pi_{t+h} = \pi_t + \varepsilon_t, \quad (8)$$

em que $\varepsilon_t \sim N(0,1)$.

b) Autorregressivo (AR)

$$\pi_{t+h} = \phi_{0,h} + \phi_{1,h}\pi_t + \dots + \phi_{p,h}\pi_{t-p+1} + \varepsilon_t \quad (9)$$

em que p é a ordem de integração dos componentes autorregressivos determinados seguindo o procedimento de Box e Jenkins (1976). A cada passo à frente um novo modelo é estimado.

c) Vetor Autorregressivo (VAR)

$$Y_t = \sum_{i=1}^p \Phi_i Y_{t-i} + \Gamma_t + v_t \quad (10)$$

no qual Y_t e v_t são vetores de dados e erros de previsão (ruído branco), ao passo que Φ_i é a matriz dos coeficientes e Γ_t contém os termos determinísticos (constante, tendência e dummies diversas). A escolha das variáveis usadas no VAR foi obtida a partir das possíveis combinações do conjunto de variáveis: Índice de preço, PIB, taxa de juros, taxa de câmbio e agregado monetário M1.

d) Vetor Autorregressivo com parâmetros variantes no tempo (TVP-VAR)

O modelo baseia-se em Primiceri (2005) que propôs o seguinte modelo para um vetor de série temporal n-dimensional Y_t :

$$Y_t = c_t + B_{1,t}Y_{t-1} + \dots + B_{p,t}Y_{t-p} + A_t^{-1}\Sigma_t\varepsilon_t \quad (11)$$

em que o intercepto c_t e a matriz de coeficientes $\{B_{j,t}\}_{j=1}^p$ são variantes no tempo. Além disso, o termo de erro $A_t^{-1}\Sigma_t\varepsilon_t$ implica que a matriz de variância-covariância dos resíduos (VCV) também pode variar ao longo do tempo. Mais precisamente, o modelo completo é dado pelas seguintes equações:

$$\begin{aligned}
Y_t &= X_t B_t + A_t^{-1} \Sigma_t \varepsilon_t \\
B_t &= B_{t-1} - \nu_t \\
\alpha_t &= \alpha_{t-1} + \zeta_t \\
\log \sigma_t &= \log \sigma_{t-1} + \eta_t
\end{aligned} \tag{12}$$

em que $X_t' = I_n \otimes [1, Y_{t-1}, \dots, Y_{t-p}]$, B_t recolhe os parâmetros c_t e $\{B_{j,t}\}_{j=1}^p$ da Equação (12), A_t é uma matriz triangular inferior (com valor 1 na diagonal principal) cujos elementos livres são empilhados no vetor α_t , e Σ_t é uma matriz diagonal com elementos positivos $\sigma_t = \text{diag}(\Sigma_t)$. ε_t segue uma distribuição normal padrão, e $\{\nu_t, \zeta_t, \eta_t\}$ são vetores aleatórios Gaussianos de média zero, homocedásticos e mutuamente independentes de dimensões apropriadas.

e) ARMA-Exponential GARCH (Generalized Autoregressive Conditional Heteroskedasticity)

O modelo exponencial de Nelson (1991)⁷ é definido como

$$\pi_{t+1} = \mu + \theta_1 \pi_t + \theta_2 \pi_{t+1} + \dots + \theta_p \pi_{t-p+1} + \phi_1 \varepsilon_t + \phi_2 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p+1} \tag{13}$$

$$\log \sigma_{t+1}^2 = w + \alpha \left| \frac{\varepsilon_t}{\sigma_t} \right| + \delta \frac{\varepsilon_t}{\sigma_t} + \beta \log \sigma_t^2 \tag{14}$$

em que $\log \sigma_t^2$ é o logaritmo da variância condicional, que fornece variância condicional positiva mesmo se os parâmetros forem negativos. O efeito assimétrico pode ser testado pela hipótese de que $\delta \neq 0$. O parâmetro β captura a persistência na volatilidade condicional. A seleção do modelo é baseada no critério de informação de Akaike (AIC), sendo o modelo estimado pelo método da verossimilhança (algoritmo BFGS⁸), permitindo até quatro defasagens. Neste trabalho serão usadas distribuições condicionais normal e *t-Student* como alternativas para ajustar as inovações padronizadas.

f) Modelo de previsão baseado na Curva de Phillips

Modelos baseados na curva de Phillips são amplamente utilizados para gerar previsões da taxa de inflação. Trabalhos como os de Stock e Watson (1996, 1999, 2007), Atkeson e Ohanian (2001), Ascari e Marrocu (2003) e Rumler e Valderrama (2010) e Arruda et al. (2011) utilizam diversas especificações da curva de Phillips para prever a inflação.

A equação abaixo representa a curva de Phillips em sua versão ampliada

$$\pi_{t+1} = \alpha_1 \pi_{t-p+1} + \alpha_2 h_{t-p+1} + \alpha_3 (\Delta e_{t-p+1} + \pi_{t-p+1}^*) + \varepsilon_t \tag{15}$$

⁷ Também foram realizadas estimações com outros modelos da família GARCH, como GJRGARCH e TARCH, mas não apresentam resultados tão bons em termos preditivos.

⁸ Na otimização numérica, o algoritmo Broyden-Fletcher-Goldfarb-Shanno (BFGS) é um método iterativo para resolver problemas de otimização não linear sem restrições.

em que h_t é o hiato do produto; o repasse cambial é medido por $\Delta e_{t-p+1} + \pi_{t-p+1}^*$, onde e_t o logaritmo da taxa de câmbio nominal, Δ é o operador diferença e π_t^* é uma medida de inflação internacional (PPI⁹ americano).

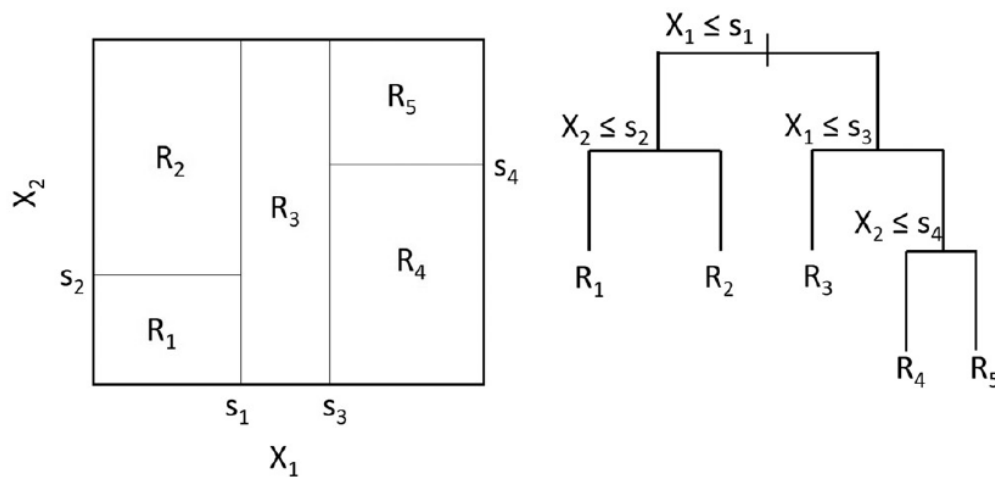
1.4.2.2. Modelos de previsão baseados em *Machine Learning*

Também foram usados modelos de previsão baseados em técnicas de *Machine Learning: Random Forest* (Breiman, 2001) e *least absolute shrinkage and selection operator* (Tibshirani, 1996). Para ambos os modelos foi aplicado o algoritmo produzido por Redell (2020) inspirado em Bergmeir et al. (2018).

a) *Machine Learning-Random Forest* (ML-RF)

O modelo ML-RF é uma técnica de aprendizado supervisionado proposto por Breiman (2001) para reduzir a variância de árvores de regressão e é baseado na agregação de bootstrap (*bagging*) de árvores de regressão construídas aleatoriamente. Por sua vez, uma árvore de regressão é um modelo não paramétrico que aproxima uma função não linear desconhecida com predições locais usando particionamento recursivo do espaço das covariáveis (Breiman, 1996). Para entender como uma árvore de regressão funciona, um exemplo de Hastie et al. (2009) é útil. Considere um problema de regressão em que X_1 e X_2 são variáveis explicativas, cada uma tomando valores em um determinado intervalo, e Y é a variável dependente. Primeiro, divide-se o espaço em duas regiões em $X_1 = s_1$, e então a região à esquerda (direita) de $X_1 = s_1$ é dividida em $X_2 = s_2$ ($X_1 = s_3$). Finalmente, a região à direita de $X_1 = s_3$ é dividida em $X_2 = s_4$. Conforme ilustrado no gráfico à esquerda da Figura 4, o resultado é uma partição em cinco regiões: $R_k, k = 1, \dots, 5$.

Figura 4 - Exemplo de uma árvore de regressão



Fonte: Figura 9.2 de Hastie et al. (2009), p.306 retirado em 2021.

⁹ *Producer Price Index*.

Em cada região R_k , assume-se que o modelo prevê π com uma constante c_k , que é estimado como a média amostral de realizações de Y que “caem” dentro da região R_k . Uma vantagem importante dessa partição binária recursiva é que ela pode ser representada como uma única árvore, conforme ilustrado no gráfico à direita da Figura 4. Cada região corresponde a um nó terminal da árvore.

Dada uma variável dependente π_{t+h} , um conjunto de preditores \mathbf{x}_t e um número de nós terminais K , as divisões são determinadas para minimizar a soma dos erros quadrados do seguinte modelo de regressão

$$\pi_{t+h} = \sum_{k=1}^K c_k l_k(\mathbf{x}_t; \boldsymbol{\theta}_k) \quad (16)$$

Em que $l_k(\mathbf{x}_t; \boldsymbol{\theta}_k)$ é a função indicadora tal que

$$l_k(\mathbf{x}_t; \boldsymbol{\theta}_k) = \begin{cases} 1 & \text{se } \mathbf{x}_t \in R_k(\boldsymbol{\theta}_k) \\ 0 & \text{caso contrário} \end{cases}$$

$\boldsymbol{\theta}_k$ é o conjunto de parâmetros que define a k -ésima região. $l_k(\mathbf{x}_t; \boldsymbol{\theta}_k)$ é um produto de funções indicadoras, cada uma das quais define uma das divisões que produz a k -ésima região. ML-RF é uma coleção de árvores de regressão, cada uma especificada em uma amostragem *bootstrap* dos dados originais. Suponha que existam B amostras *bootstrap*. Para cada amostra b , $b = 1, \dots, B$, uma árvore com regiões K_b é estimada para um subconjunto selecionado aleatoriamente dos regressores originais. K_b é determinado para deixar um número mínimo de observações em cada região. A previsão final é a média das previsões de cada árvore aplicada aos dados originais

$$\hat{\pi}_{t+h} = \frac{1}{B} \sum_{b=1}^B \left[\sum_{k=1}^{K_b} \hat{c}_{k,b} l_{k,b}(\mathbf{x}_t; \hat{\boldsymbol{\theta}}_{k,b}) \right]$$

b) ML-LASSO

O LASSO foi inicialmente proposto por Tibshirani (1996). A ideia por trás do método é reduzir coeficientes irrelevantes em uma regressão para zero. Isso é feito adicionando uma penalidade à função de perda diretamente os parâmetros estimados, excluindo aqueles que são irrelevantes. O estimador LASSO é definido como:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} \left[\sum_{t=1}^T (\pi_{t+h} - \boldsymbol{\beta}' \mathbf{x}_t)^2 + \lambda \sum_{j=1}^q |\beta_j| \right] \quad (17)$$

em que $\boldsymbol{\beta} \in \mathbb{R}^q$ um vetor de parâmetros desconhecidos, $\mathbf{x}_t = (x_{1t}, \dots, x_{qt})' \in \mathbb{X} \subseteq \mathbb{R}^q$ é a matriz de q preditores e λ é o parâmetro *shrinkage* (encolhimento).

Se for definido $\hat{\boldsymbol{\beta}}_0$ como o estimador de mínimos quadrados ordinários (OLS) sem qualquer encolhimento e λ_{ols} como o maior parâmetro de penalidade que fornece o mesmo resultado que o OLS (sem encolhimento), haverá encolhimento para qualquer $\lambda > \lambda_{ols}$. Além disso, seja λ_c o

menor parâmetro de penalidade que resulta em um modelo apenas com um termo constante. Então, deve-se escolher no intervalo $\lambda_{ols} \leq \lambda \leq \lambda_c$. O melhor é escolhido usando validação cruzada ou algum critério de informação como o *Bayesian Information Criterion* (BIC). Seguindo Medeiros et. al (2016), a equação estimada é definida em

$$\pi_{t+h} = \alpha_0 + \sum_{i=0}^3 \gamma_i \pi_{t-i} + \sum_{i=0}^3 \beta' x_{t-i} + u_{t+h} \quad (18)$$

onde α_0 é a constante, x_t é o vetor que contém todas as variáveis candidatas e u_t é o termo de erro. A equação (18) afirma que a inflação mensal h períodos à frente é função da inflação de hoje e dos 3 meses anteriores e de outras variáveis de controle.

1.5. Dados

A inflação foi medida pelo Índice de Preços ao Consumidor Brasileiro (IPCA), que é o índice oficial de inflação no Brasil. Além disso, um número considerável de títulos atrelados à inflação usa o IPCA como referência. A taxa de inflação mensal, π , utilizada foi obtida da seguinte forma: $\pi_t = \left(\left(\frac{ipca_t}{ipca_{t-1}} \right) - 1 \right) \times 100$.

Todos os dados utilizados neste estudo compreendem o período de janeiro de 1996 a dezembro de 2019 com frequência mensal, um total de 288 observações.

As variáveis da Tabela 2 são as mais importantes na teoria econômica que afetam o nível geral de preços, e que foram utilizadas nos modelos econométricos apresentados na Seção anterior. Além das variáveis de índice de preço apresentadas na Seção 1.1. A amostra foi dividida em dois períodos: (i) período de treinamento (dentro da amostra) - janeiro de 1996 a dezembro de 2014; (ii) período de teste (fora da amostra) – janeiro de 2015 a dezembro de 2019.

Tabela 2 - Banco de Dados (continua)

Sigla	Descrição	Sigla	Descrição
PIB	Crescimento real do produto interno bruto do Brasil	Horas	Horas trabalhadas na indústria brasileira
hiato	Hiato do produto interno bruto do Brasil	Onibus	Tarifa de Onibus em SP
IBOV	Retornos mensais do Ibovespa	Energia	Tarifa de Energia
Petr	Retornos do preço do Petróleo	IMA.B	Índice de Mercado Anbima dos títulos públicos federais atrelados ao IPCA
M0	Base monetária	IMA.B.5-	Índice de Mercado Anbima dos títulos públicos federais atrelados ao IPCA com maturidade inferior a cinco anos

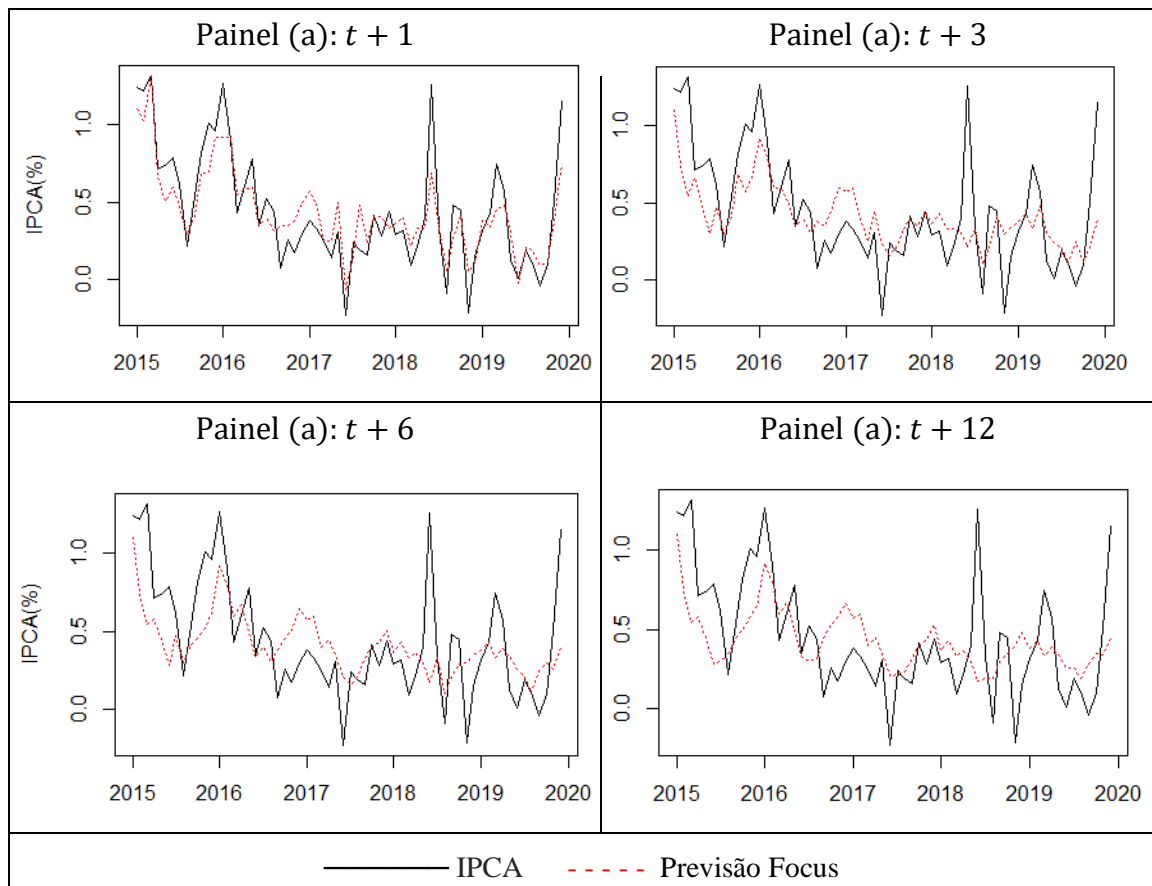
Tabela 2 - Banco de Dados (conclusão)

M1	Agregado monetário M1	IMA.B.5+	Índice de Mercado Anbima dos títulos públicos federais atrelados ao IPCA com maturidade superior a cinco anos
M2	Agregado monetário M2	Sup.Prim	Superavit Primário
M3	Agregado monetário M3	DSLPe	Varição da dívida líquida do setor Público – Estados brasileiros
M4	Agregado monetário M4	DSLPe	Varição da dívida líquida do setor Público – Governo federal brasileiro
Over Selic	Taxa de juros praticada no mercado	e-dolar	Taxa de câmbio USD-BRL
Selic.copom	Taxa de juros fixada pelo Comitê de Política Monetária	e-turism	Taxa de câmbio turística USD-BRL
Desemp	Taxa de desemprego do Estado de São Paulo	e-euro	Taxa de câmbio EUR-BRL

Fonte: O Autor, 2022

O banco de dados total é composto por 9 variáveis relacionados a índices de preços e 24 relacionados a outras variáveis macroeconômicas, totalizando 33 variáveis. A fontes dos dados são: Bloomberg, Banco Central do Brasil (BCB) e Instituto Brasileiro de Geografia e Estatística (IBGE).

Figura 5 - Índice de Preço ao Consumidor Amplo (IPCA) e previsões do Boletim Focus.



Fonte: O Autor, 2022. Elaboração própria com base nos dados do IBGE (IPCA) e Banco Central (Previsão Focus)

Para verificar a estacionariedade dos dados foram aplicados os testes de Dickey Fuller Aumentado (ADF), Phillips Peron (PP) e Kwiatkowski – Phillips – Schmidt – Shin (KPSS) para todas as séries de dados mencionadas acima, em nível e em variação mensal. Os resultados estão no Anexo.

O conjunto de dados também inclui previsões de especialistas da pesquisa Focus produzida pelo BCB. Nossas variáveis de expectativa incluem a mediana das previsões do especialista h -períodos à frente e a mediana dos cinco maiores (Top5Focus) especialistas, ou seja, os cinco especialistas que produziram as melhores previsões no período anterior.

Tanto a série do IPCA quanto a mediana das previsões do Boletim Focus são apresentadas na Figura 5. Observe que as previsões do Focus oferecem pequenos erros de previsão para $h = 1$, mas perde rapidamente o desempenho conforme h aumenta.

1.6. Seleção de variáveis

Em economia, existe uma disponibilidade muito grande de dados que podem ser utilizados como variáveis preditoras em um modelo de previsão. Diante disso, quando o número de covariáveis é muito grande é necessário selecionar quais as mais importantes que vão compor o modelo. Os principais métodos para a seleção de variáveis são: seleção do melhor subconjunto de covariáveis, *backward*, *forward*, *stepwise*, rede elástica (*elastic net*) e LASSO.

Ao estimar parâmetros em grandes dimensões, os métodos de *Shrinkage* (encolhimento) constituem uma alternativa bem-sucedida. A ideia é reduzir a zero os parâmetros que correspondem às variáveis irrelevantes. Sob algumas condições, é possível lidar até mesmo com mais variáveis do que observações.

Entre os métodos de encolhimento, a regressão linear rede elástica se mostra uma boa alternativa, utilizando tanto a penalidade das técnicas de regressão LASSO quanto da regressão em crista (*ridge*) para selecionar e regularizar os modelos.

Este método de seleção e redução de covariáveis foi implementado por muitos autores. Os principais suportes teóricos são: Tibshirani (1996), Zou et al. (2007), Wang et al. (2007), Friedman et al. (2010), Murphy (2012), Hastie et al. (2015) e Casagrande (2016).

Para evitar confusão, é importante ressaltar que este método de seleção de variáveis é uma etapa em comum ao modelo ML-LASSO. Ou seja, esta técnica também foi aplicada no ML-LASSO, porém para realizar a previsão da inflação ainda é necessário seguir o procedimento adicional apresentado na Subseção 5.2.2. Já nesta Seção, apenas foi aplicada a técnica para selecionar as melhores variáveis preditoras da inflação.

1.6.1. Regressão linear de rede elástica e LASSO

Suponha o vetor de observações (variáveis preditoras) $\mathbf{X}_i \in \mathbb{R}$ e a variável resposta $\pi_i \in \mathbb{R}, i = 1, \dots, N$. A função objetivo da regressão linear de rede elástica é

$$\min_{(\beta_0, \beta) \in \mathbb{R}} \frac{1}{2N} \sum_{i=1}^N (\pi_i - \beta_0 - \mathbf{X}_i' \beta)^2 + \lambda \left[\frac{(1-\alpha) \|\beta\|_2^2}{2} + \alpha \|\beta\|_1 \right], \quad (19)$$

sendo β_0 o intercepto e $\lambda \geq 0$ é um parâmetro que controla o nível de encolhimento e é determinado por técnicas baseadas em dados, como validação cruzada ou o uso de critérios de informação (será discutida a escolha adequada de λ mais à frente).

O termo $\lambda \left[\frac{(1-\alpha) \|\beta\|_2^2}{2} + \alpha \|\beta\|_1 \right]$ representa a penalização sugerida, se $\alpha = 1$, será a penalização do tipo L_1 , correspondente à regressão LASSO, e se $\alpha = 0$, a penalização será do tipo L_2 , correspondente à regressão em crista. Note que se $0 < \alpha < 1$, tanto a penalização L_1 quanto L_2 serão incluídas, neste caso é chamado de rede elástica.

O caso específico do estimador via LASSO é basicamente um problema de minimização usual dos erros ao quadrado acrescido de uma penalização L_1 em relação aos parâmetros de posição do modelo de regressão. Nesse contexto, a resolução do problema será

$$\hat{\beta}_L = \arg \min_{\beta} \left(\sum_{i=1}^n (\pi_i - \mathbf{X}_i \beta)^2 + \lambda \sum_{j=1}^p |\beta_j| \right), \quad \lambda \geq 0, \quad (20)$$

Um detalhe importante acerca da penalidade é que o parâmetro β_0 não é restringido pela penalização. Note que quanto maior for o valor de λ maior será a penalização.

Segundo Hastie et al. (2015), o estimador LASSO pode não ter a performance esperada quando as variáveis são muito correlacionadas. Adicionar o componente $\frac{1}{2}(1-\alpha) \|\beta\|_2^2$ na penalização ajuda a controlar fortes correlações entre grupos de variáveis, mantendo a característica de tornar modelos esparsos que o LASSO possui. Em contrapartida, temos mais um parâmetro a estimar, α , que determina a proporção de cada penalização, em crista e LASSO, que será aplicada.

Assim, os parâmetros λ e α são responsáveis por controlar a força da penalidade e construir uma ponte entre a regressão em crista e o LASSO, respectivamente (Friedman et al., 2010). O primeiro é iterado quantas vezes se desejar, já o segundo é determinado pelo analista, variando no intervalo $[0, 1]$.

Sob essa perspectiva, neste estudo, a seleção de variáveis será feita pela regressão linear de rede elástica e pelo LASSO, seu caso particular.

1.6.2. Escolhendo lambda

Segundo Tibshirani (1996), para que o método funcione adequadamente, o parâmetro de ajuste λ desempenha um papel fundamental, pois controla a quantidade de “encolhimento” aplicada às estimativas. Quando maior o encolhimento, menor o número de covariáveis preditoras.

Um modo de escolher o valor de λ é a validação cruzada, em que a amostra é dividida aleatoriamente em k partes iguais ou pelo menos aproximadamente iguais. A técnica consiste em, a cada rodada, separar um grupo para ser o grupo de teste (*test set*) e os demais $k - 1$ compõem o grupo de treinamento (*training set*), de modo que esse procedimento se repita k vezes e todos os grupos sejam o grupo de teste uma vez. Após terminar as k iterações, tem-se k erros de predição e será calculada a média dos k erros de predição. Esse procedimento será feito para vários valores λ distintos. Finalmente, seleciona-se o valor de λ que minimize o erro de predição médio.

Tibshirani (1996) recomenda $k = 5$ ou $k = 10$, porém não há uma regra para escolha de k . Claro que o número de observações deve ser “suficientemente grande” para ser dividido entre as k partes. Para cada rodada da validação cruzada o grupo de treinamento será usado para estimar os coeficientes e ajustar modelos para uma gama de valores de λ . Em seguida, aplica-se os valores do grupo de teste nesta gama de modelos e compara-se os valores reais do grupo com os valores preditos. Os valores de λ aplicados são arbitrários e o ideal é que sejam muitos.

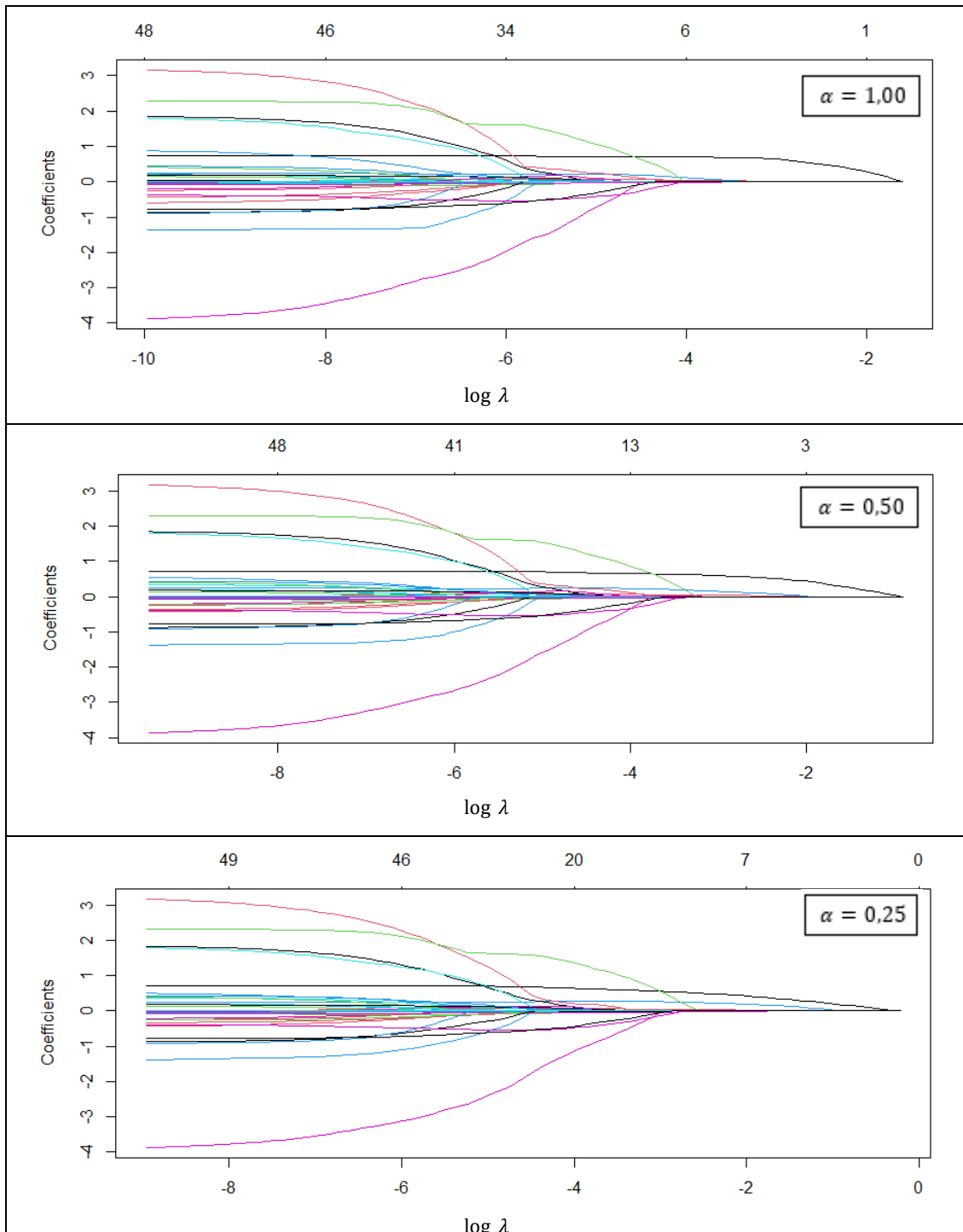
Ao final, teremos um conjunto de erros de predição (EP) e erros quadráticos médios (EQM) para cada valor de λ . A medida utilizada para determinar o parâmetro de ajuste ótimo costuma ser o EQM, uma vez que seu cálculo é computacionalmente mais simples do que o do EP (Casagrande, 2016). Assim, calcula-se a média dos valores de EQM para cada λ a fim de escolher o que apresente a menor média dos erros quadráticos médios.

Para este trabalho, foram utilizados dois valores de λ : (i) λ_{min} , o valor de λ que minimiza o erro médio da validação cruzada, e (ii) λ_{1se} , valor de λ que fornece o modelo mais regularizado, de modo que o erro de validação cruzada está dentro de um desvio padrão do mínimo.

Antes de estimar os modelos de previsão da inflação, será realizada uma seleção das melhores variáveis preditoras ajustando o método de regressão linear de rede elástica aos dados. Serão testadas 27 variáveis, além de algumas defasadas, totalizando 49 variáveis potenciais preditoras.

A Figura 6 apresenta os gráficos das estimativas dos parâmetros via regressão linear de rede elástica. Cada curva representa uma variável preditora e o eixo horizontal superior indica o número de coeficientes diferentes de zero para o valor de λ correspondente do eixo horizontal inferior.

Figura 6 - Gráfico da estimativa dos parâmetros via regressão linear de rede elástica



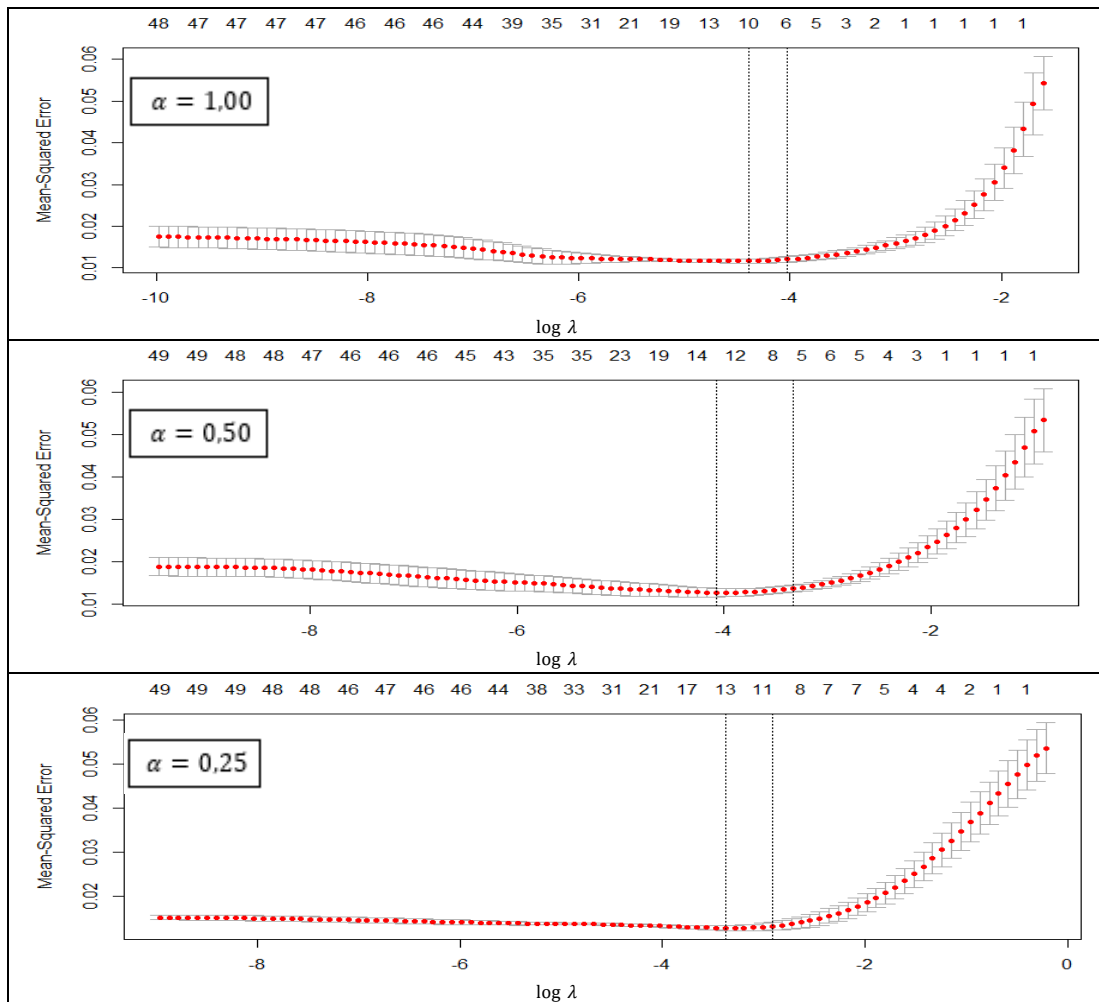
Fonte: O Autor, 2022.

Observe que à medida que o $\log \lambda$ aumenta, os coeficientes de cada variável preditora diminuem até se igualarem a zero. Quanto menores os valores de α , maior o número de variáveis preditoras “sobreviventes”. Nota-se o poder de “encolhimento” que o LASSO ($\alpha = 1$) possui, por exemplo, se $\log \lambda = -4$, somente em torno de 6 variáveis preditoras teriam coeficientes diferentes de zero, o que é equivalente a dizer que apenas estas foram selecionadas.

A escolha do valor de λ é através do método de validação cruzada. Seguindo Tibshirani (1996), a amostra foi dividida em 10 partes iguais, isso representa sub amostras de 24 meses, e assim aplicada a validação cruzada dos dados com base no critério do erro quadrático médio.

A Figura 7 apresenta as curvas de validação cruzada (linha pontilhada vermelha) junto com as curvas de desvio padrão superior e inferior ao longo do eixo $\log \lambda$ (barras de erro). Este gráfico é muito elucidativo, pois mostra o comportamento do erro quadrático médio ao longo das iterações, auxiliando na escolha do valor de λ . Dois valores especiais de λ estão indicados pelas linhas pontilhadas verticais: λ_{min} é o valor de λ que minimiza o erro de validação cruzado médio, enquanto λ_{1se} é o valor de λ que fornece o modelo mais regularizado, de modo que o erro de validação cruzada está dentro de um desvio padrão do mínimo. Já os valores do eixo horizontal superior representam o número de variáveis explicativas com o coeficiente diferente de zero. A Tabela 3 apresentam os coeficientes de cada modelo ajustado via regressão linear de rede elástica para diferentes alfas e para λ_{min} e λ_{1se} .

Figura 7 - Curva de Validação Cruzada



Fonte: O Autor, 2022

Nota: Curva de validação cruzada para regressão linear de rede elástica diferentes valores de alfa.

Tabela 3 - Estimativas dos parâmetros do modelo ajustado via Regressão linear de rede elástica para λ_{min} e λ_{1se}

	$\alpha = 1,00$	$\alpha = 1,00$	$\alpha = 0,50$	$\alpha = 0,50$	$\alpha = 0,25$	$\alpha = 0,25$
($\times 10$)	λ_{min}	λ_{1se}	λ_{min}	λ_{1se}	λ_{min}	λ_{1se}
Intercept	1.172	0.870	1.040	0.996	1.535	1.246
<i>IPCA-15_t</i>	5.490	5.836	5.390	5.503	4.332	5.068
<i>IGP-DI_t</i>	0.191	0.360	0.363	0.386	0.221	0.324
<i>IGP-M_t</i>	0.160	0.124
<i>IPC-Br_t</i>	1.830	1.945	2.011	2.008	1.976	2.059
<i>IGP-10_t</i>
<i>selic_t</i>
<i>PIB_t</i>
<i>hiato_t</i>	0.020
<i>desemp_t</i>	.	.	.	-0.017	.	-0.049
<i>petro_t</i>
<i>e-dolar_t</i>
<i>e-turism_t</i>
<i>e-euro_t</i>
<i>nucleo_t</i>
<i>IBOV_t</i>
<i>DSLPe_t</i>
<i>DSLpb_t</i>
<i>IMA-B_t</i>
<i>IMA-B5+_t</i>
<i>IMA-B.5-_t</i>
<i>M0_t</i>
<i>M1_t</i>	0.196
<i>M2_t</i>
<i>M3_t</i>
<i>M4_t</i>	.	.	.	-2.603	.	-7.568
<i>horas_t</i>	-0.001
<i>onibus_t</i>
<i>energia_t</i>
<i>selic_{t-1}</i>
<i>PIB_{t-1}</i>
<i>desemp_{t-1}</i>	.	-0.018	-0.025	-0.052	-0.015	-0.103
<i>petro_{t-1}</i>
<i>e-dolar_{t-1}</i>
<i>e-turism_{t-1}</i>
<i>e-euro_{t-1}</i>
<i>nucleo_{t-1}</i>
<i>IBOV_{t-1}</i>	0.003	0.023	0.020	0.024	0.010	0.025
<i>DSLPe_{t-1}</i>
<i>DSLpb_{t-1}</i>
<i>IMA-B_{t-1}</i>
<i>IMA-B5+_{t-1}</i>
<i>IMA-B.5-_{t-1}</i>
<i>M0_{t-1}</i>
<i>M1_{t-1}</i>
<i>M2_{t-1}</i>
<i>M3_{t-1}</i>	.	.	.	1.016	.	2.492
<i>M4_{t-1}</i>
<i>horas_{t-1}</i>
<i>onibus_{t-1}</i>	.	-0.867	-0.068	-0.900	.	-1.013
<i>energia_{t-1}</i>

Fonte: O autor, 2022.

Apenas algumas variáveis, além do intercepto, foram selecionadas. Existe uma pequena diferença entre o número de variáveis selecionadas para cada valor de alfa. Todas as variáveis

com coeficientes diferentes de zero serão testadas como exógenas aos modelos GAS. O objetivo é incorporar um grupo de variáveis que aumentem o poder preditivo dos modelos.

1.7. Avaliação da previsão

Nesta Seção, será avaliado o desempenho de previsão dos modelos. As várias especificações do GAS são avaliadas em relação aos modelos competidores para previsão de inflação. A aplicação empírica do exercício de previsão para 1, 3, 6 e 12 meses à frente é realizada em um esquema de janela móvel de cinco anos (60 meses), com cada modelo sendo reestimado recursivamente em cada ponto de previsão. Por exemplo, para previsão de um passo à frente, primeiramente os parâmetros do modelo são estimados para o período dentro da amostra e prevista a inflação do mês seguinte fora da amostra. Como a estimação é feita em um esquema de janela móvel, a informação mais antiga é descartada e a informação observada mais recente é incorporada formando um novo período dentro da amostra, e assim realizada uma nova estimação dos parâmetros do modelo e a previsão do mês seguinte fora da amostra. Esse procedimento é repetido até completar os 60 meses de previsão fora da amostra.

Relembrando que a amostra foi dividida em dois períodos: (i) período de treinamento (dentro da amostra) - janeiro de 1996 a dezembro de 2014; (ii) período de teste (fora da amostra) – janeiro de 2015 a dezembro de 2019. A Tabela 4 relata as abreviações de cada modelo com sua respectiva especificação.

Tabela 4 - Abreviações usadas no estudo para fins de representação

Abreviação	Modelo	Abreviação	Modelo
GAS-Ng	GAS Normal e homocedástico	GARCH-N	ARMA - exponencial normal GARCH
GAS-tg	GAS <i>t-Student</i> e homocedástico	GARCH-t	ARMA - exponencial Student-t GARCH
GAS-N	GAS Normal e heterocedástico	VAR	Vetor autorregressivo
GAS-t	GAS <i>t-Student</i> e heterocedástico	TVP VAR	Vetor autorregressivo com parâmetros variante no tempo
RW	Random Walk	Phillips	Curva de Phillips - versão ampliada
AR-N	AR normal	ML-RF	Machine learning- Random forest
AR-t	AR <i>t-Student</i>	ML-LASSO	Machine learning- LASSO

Fonte: Fonte: O autor, 2022.

Nota: Esta tabela pode ser um ponto de referência para a leitura das demais tabelas nas quais os resultados dos modelos são apresentados

Foram estimadas quatro especificações de modelos GAS: Distribuição Normal ou *t-Student* e homocedástico ou heterocedástico. Para cada uma foram incorporadas variáveis preditores selecionadas na Seção anterior.

A análise do desempenho das previsões foi conduzida a partir de perspectivas distintas. Primeiro, avaliação foi realizada através previsão pontual via erros de previsão e desempenho na presença de instabilidades. Posteriormente, em termos de densidade preditiva.

1.7.1. Previsão pontual

Esta Subseção discute os resultados e compara os erros de previsão de todos os modelos. A Tabela 5 mostra a raiz do erro quadrático médio (EQM) e o erro absoluto médio (EAM) para todos os modelos. O modelo estimado neste estudo com o menor erro de previsão para cada horizonte é exibido em **negrito**.

O Focus e Top5Focus têm os menores erros para todos os horizontes de previsão. No entanto, para horizontes mais longos, a diferença de erros entre estes e os modelos estimados se reduz, tornando a comparação mais competitiva. Os modelos GAS foram os que tiveram melhor performance quando comparados aos outros modelos estimados neste estudo, com destaque para GAS-t (em negrito), que apresentou os melhores resultados para 1, 3, 6 e 12 passos à frente. Observa-se que o uso da distribuição *t-Student* e a inclusão da heteroscedasticidade condicional são importantes para melhorar o desempenho.

Após verificar a performance individual dos modelos combinados, utilizou-se o teste de Giacomini-White (Giacomini e White, 2006) para realizar uma análise comparativa entre os métodos de previsão.

Tabela 5 - Erros quadráticos médios (EQM) e erros absolutos médios (EAM) de previsão

Horizonte de Previsão	$t + 1$		$t + 3$		$t + 6$		$t + 12$	
	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM
RW	3.28	2.54	3.73	2.87	4.04	3.42	4.66	3.61
AR-N	2.73	1.82	3.22	2.45	3.82	2.85	3.96	2.86
AR-t	2.37	1.62	3.21	2.38	3.42	2.78	3.64	3.4
GARCH-N	2.56	1.85	3.15	2.37	3.56	2.92	3.81	2.95
GARCH-t	2.34	1.69	3.19	2.39	3.43	2.83	3.54	2.74
GAS-Ng	2.30	1.56	3.29	2.53	3.29	2.69	3.65	2.92
GAS-tg	2.28	1.52	2.98	2.32	3.37	2.73	3.32	2.56
GAS-N	2.28	1.50	2.94	2.18	3.32	2.73	3.53	2.8
GAS-t	2.21	1.48	2.98	2.31	3.29	2.73	3.31	2.61
VAR	2.37	1.71	3.06	2.27	3.53	2.88	3.59	2.84
TVP VAR	2.33	1.66	3.01	2.27	3.29	2.75	3.51	2.64
Phillips	2.52	1.82	3.04	2.25	3.53	2.89	3.55	2.89
ML-RF	2.48	1.80	3.25	2.35	3.53	2.85	3.81	2.79
ML-LASSO	2.44	1.79	3.04	2.16	3.44	2.68	3.58	2.68
Focus	1.72	1.34	2.90	2.07	3.06	2.32	3.14	2.44
Top5Focus	1.69	1.26	2.85	1.99	3.05	2.29	3.13	2.43

Fonte: O Autor, 2022.

Nota: (EQM \times 10) – Raiz do erro quadrático médio; (EAM \times 10) - Erro absoluto médio. Os modelos indicados em negrito são os melhores de acordo com cada medida de erro e para cada horizonte de previsão. A área sombreada mostra as previsões do Boletim Focus.

Essa abordagem permite considerar os efeitos da incerteza da estimação na performance das previsões. O teste funciona da seguinte forma: sejam dois modelos $f_t(\beta_1)$ e $g_t(\beta_2)$ para h passos à frente da variável π_{t+h} , usando a função de perda $L_{t+h}(\cdot)$. A função de perda depende dos parâmetros estimados $\hat{\beta}_{1t}$ e $\hat{\beta}_{2t}$, e a função expectativa condicional depende de algum conjunto de informação G_t :

$$H_0: E \left[L_{t+h} \left(\pi_{t+h}, f_t(\hat{\beta}_{1t}) \right) - L_{t+h} \left(\pi_{t+h}, g_t(\hat{\beta}_{2t}) \right) | G_t \right] = 0 \quad (21)$$

Deste modo, a hipótese nula se torna $E [\Delta L_{t+h} | G_t] = 0$ para $t = 1, 2; \dots$. Sua hipótese nula diz que não é possível dizer qual modelo é mais preciso na previsão $t + h$ usando as informações em G_t .

Elaborando a primeira previsão de h passos à frente no instante m , utilizando o procedimento de estimação por janela móvel, Giacomini e White (2006) mostram que dado um conjunto q de funções mensuráveis H_t , a hipótese nula de igualdade de habilidade de previsão condicional para um par de modelos M_1 e M_2 , pode ser testada usando a estatística:

$$\begin{aligned} T_{m,n,h}^H &= n \left(\frac{1}{n} \sum_{t=m}^{T-h} H_t \Delta L_{m,t+h} \right)' \tilde{\Omega}_n^{-1} \left(\frac{1}{n} \sum_{t=m}^{T-h} H_t \Delta L_{m,t+h} \right) \\ &= n (\bar{Z}'_{m,n}) \tilde{\Omega}_n^{-1} (\bar{Z}_{m,n}), \end{aligned} \quad (22)$$

em que H_t é uma função de teste $q \times 1$ mensurável, $\bar{Z}_{m,n} \equiv n^{-1} \sum_{t=m}^{T-h} Z_{m,t+h}$ onde $Z_{m,t+h} \equiv H_t \Delta L_{m,t+h}$, e

$$\begin{aligned} \tilde{\Omega}_n &= n^{-1} \sum_{t=m}^{T-h} Z_{m,t+h} Z'_{m,t+h} + n^{-1} \sum_{j=1}^{T-h} \omega_{n,j} \times \sum_{t=m+j}^{T-h} [Z_{m,t+h} Z'_{m,t+h-j} \\ &\quad + Z_{m,t+h-j} Z'_{m,t+h}] \end{aligned} \quad (23)$$

em que $\omega_{n,j}$ é a função de pesos tal que $\omega_{n,j} \rightarrow 1$ como $n \rightarrow \infty$ para cada $j = 1, \dots, h - 1$.

No caso especial de $h = 1$, $\tilde{\Omega}_n$ é simplificado para $\tilde{\Omega}_n \equiv n^{-1} \sum_{t=m}^{T-h} Z_{m,t+1} \times Z'_{m,t+1}$.

O teste a hipótese nula de igual capacidade preditiva condicional sempre que $T_{m,n,h}^H > \chi_{q,1-\alpha}^2$, onde $\chi_{q,1-\alpha}^2$ é o $(1 - \alpha)$ quantil da distribuição χ_q^2 e q são os graus de liberdade.

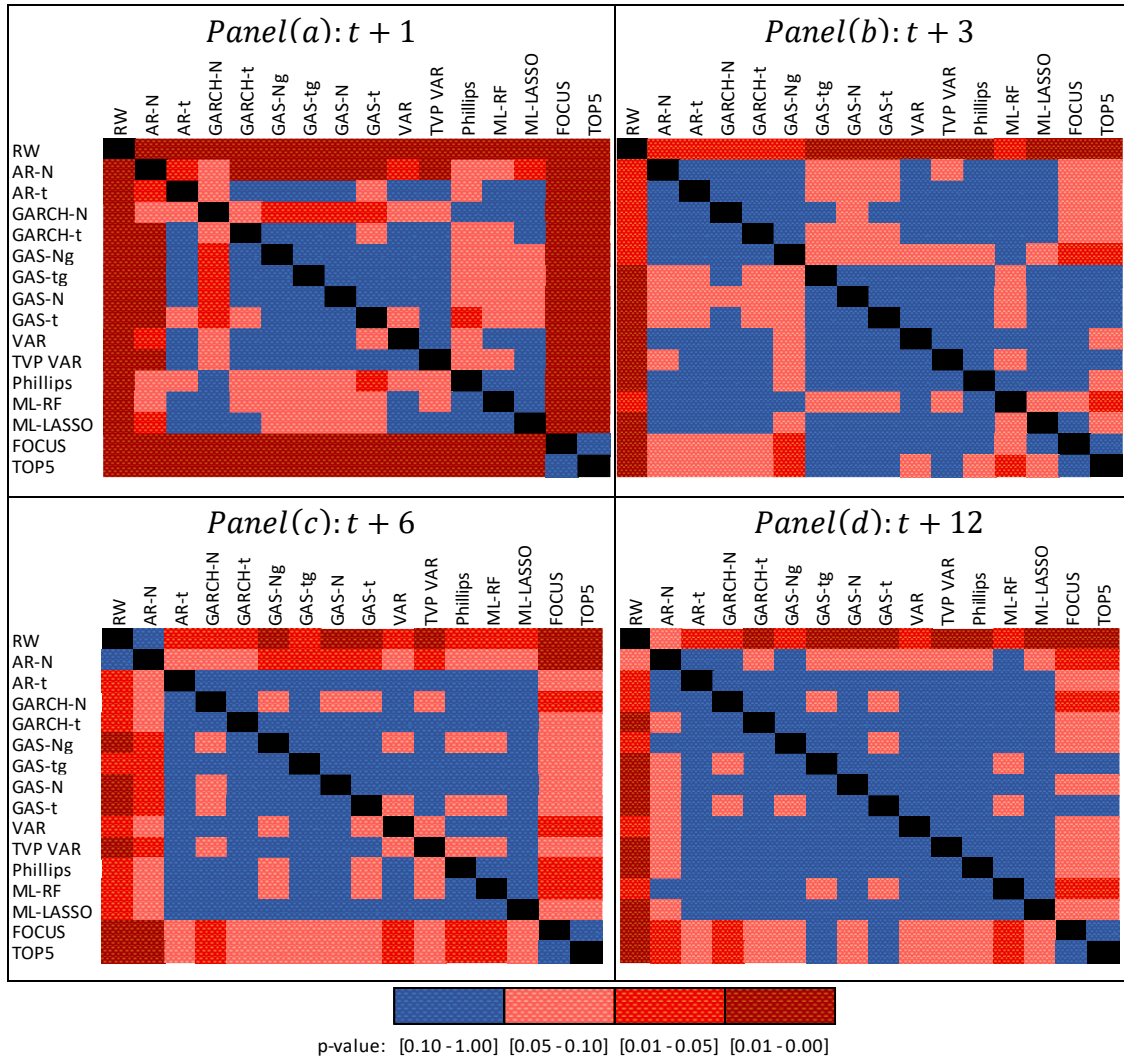
Conforme em Fernandes et al. (2014), ajustou-se o teste apresentado em Giacomini e White (2006) pelos erros padrão de Newey-West. O método de Giacomini e White (2006) foi preferido ao invés do teste de Diebold e Mariano (Diebold e Mariano, 1995), comumente utilizado, devido ao teste ser robusto à incerteza de parâmetros dos modelos, e ser aplicável a diversos métodos de estimação seja paramétrico, semi-paramétrico e não paramétrico, além de possibilitar a comparação quando técnicas de combinação de previsões são utilizadas. A Figura 8 mostra os p-valores do teste GW, lembrando que a hipótese nula é que os dois modelos possuem a mesma precisão de previsão.

O gráfico resultante é um mapa de calor em forma de tabuleiro de xadrez. O teste foi realizado em pares, com o p-valor da hipótese nula da previsão no eixo y tendo essencialmente a mesma precisão que a previsão no eixo x. Em outras palavras, p-valores próximos a 0 representam casos em que a previsão no eixo x rejeitou a hipótese da mesma precisão que a previsão no eixo y. As diferenças de erro tendem a ser menores em horizontes mais longos.

Observe como as previsões do Focus e Top5Focus rejeitaram a hipótese nula de igualdade preditiva em comparação a grande maioria dos modelos. Indicando forte

superioridade. Além disso, as previsões $t + 1$ rejeitam H_0 ao nível de significância a 1% contra qualquer modelo.

Figura 8 - Mapa de calor dos p-valores do teste GW

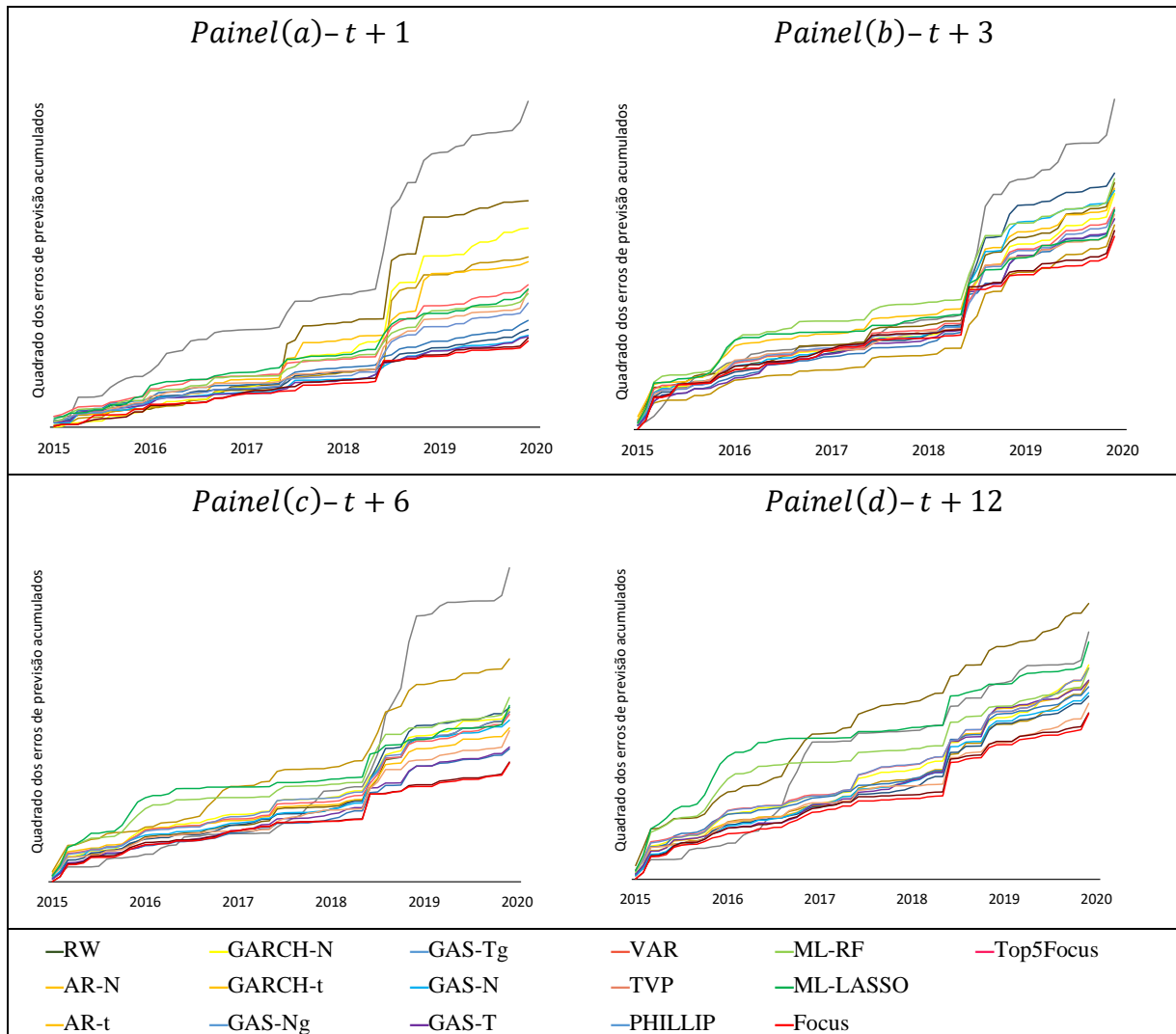


Fonte: O autor, 2022.

A Figura 9 reporta os erros quadráticos médios cumulativos para os mesmos horizontes de tempo da Tabela 5. Embora alguns modelos tenham erros menores, suas trajetórias têm comportamento semelhante.

Observe que tanto o Focus quanto o Top5Focus, representados pelas linhas vermelhas, apresentaram erros quadráticos cumulativos menores do que os outros modelos. De particular relevância é o desempenho superior em comparação aos outros para a previsão de 1 passo à frente. No entanto, para a previsão de períodos mais longos, como já verificado, Focus e Top5Focus tiveram uma queda acentuada no desempenho em relação aos demais. O motivo da mudança brusca de patamar em 2018 se refere à greve dos caminhoneiros, que reduziu temporariamente a oferta de bens e insumos, aumentando a inflação. Esse choque inflacionário aumentou abruptamente o erro de previsão em todos os modelos.

Figura 9 - Erros quadrados cumulativos: IPCA



Fonte: O Autor, 2022.

Para entender melhor o padrão e instabilidades dos resultados obtidos e explorar em que grau o desempenho preditivo relativo e absoluto variam ao longo do período fora da amostra, foram aplicados os testes de flutuação sugeridos por Giacomini e Rossi (2010) e Rossi e Sekhposian (2016). Os autores argumentaram que a média condicional e o desempenho preditivo relativo entre os modelos podem variar substancialmente durante o período fora da amostra e a média desses movimentos durante o período fora da amostra para calcular EAM ou MSE pode resultar em perda de informações. Essa questão será explorada na próxima Subseção.

1.7.2. Avaliação de previsão na presença de instabilidades

Após uma análise de capacidade preditiva, será realizada uma análise da dinâmica de racionalidade e desempenho relativo de previsão. Para isso, aplica-se os testes de desempenho (relativo) de previsão e racionalidade de previsão propostos por Giacomini e Rossi (2010) e

Rossi e Sekhposyan (2016), respectivamente, para ambientes instáveis. Esses são extremamente úteis porque permitem a inspeção visual do desempenho dos modelos de previsão ao longo do tempo.

Esses testes são usados para identificar possíveis mudanças nos desempenhos de previsão de modelos concorrentes em diferentes pontos da amostra de previsão. Como as propriedades dinâmicas dos dados macroeconômicos variam com o tempo, parece lógico suspeitar que haveria variações na capacidade dos modelos de prever a evolução futura das séries macroeconômicas. Os estudos de Rossi e Sekhposyan (2016) e Giacomini e Rossi (2010) fornecem evidências empíricas para apoiar essa visão. A seguir serão apresentadas a descrição e resultados desses dois testes, seguindo a abordagem de Rossi e Soupre (2017).

1.7.2.1. Teste de racionalidade de flutuação de Rossi e Sekhposyan (2016)

Os testes de racionalidade de previsão avaliam se as previsões satisfazem alguns requisitos “mínimos”, como ser um preditor imparcial ou não estar correlacionado com qualquer informação adicional disponível no momento da previsão. Assim, testes tradicionais de racionalidade de previsão, como os de Mincer e Zarnowitz (1969) e West e McCracken (1998), verificam se os erros de previsão têm média zero ou não estão correlacionados com alguma outra variável conhecida no momento da previsão. No entanto, eles assumem estacionaridade e, portanto, são inválidos na presença de instabilidades.

Para tornar os testes robustos a instabilidades, Rossi e Sekhposyan (2016) propuseram estimar as seguintes regressões de racionalidade de previsão em janelas móveis (de tamanho m):

$$v_{j,h} = g_j' \theta + \eta_{j,h} \quad j = t - m + 1, \dots, t \quad (24)$$

onde o erro de previsão denotado por $v_{j,h}$, refere-se a uma previsão fora da amostra h -passos à frente no instante de tempo j usando dados disponíveis até este ponto no tempo, g_j é uma função vetorial ($l \times 1$) dos dados do período j , θ é um vetor de parâmetros ($\ell \times 1$) e $\eta_{j,h}$ é o resíduo.

O principal interesse de Rossi e Sekhposyan (2016) foi testar a racionalidade da previsão na presença de instabilidades. De fato, na presença de instabilidades, os testes que focam no desempenho médio fora da amostra de um modelo podem ser enganosos porque podem calcular a média das instabilidades. Assim, as hipóteses a testar são:

$H_0: \theta_t = \theta_0$ versus $H_A: \theta_t \neq \theta_0, \forall t$ onde $\theta_0 = 0$ and θ_t é o valor do parâmetro variante no tempo.

Assim, a hipótese nula é que a previsão em consideração é racional em todos os pontos no tempo durante o período amostral. A rejeição não implica que uma previsão seja permanentemente irracional, mas que foi irracional pelo menos uma vez durante o período amostral.

Em (24), foca-se em testes de imparcialidade da previsão ($g_t = 1$), eficiência de previsão (g_t é a própria previsão) e racionalidade da previsão (g_t inclui a previsão e 1).

Refere-se aos testes sob a hipótese mantida de que $\theta_0 = 0$ como “testes de racionalidade de previsão”. A restrição zero sob a hipótese nula garante que os erros de previsão sejam

realmente imprevisíveis, dado o conjunto de dados disponível no momento em que a previsão é feita.

Rossi e Sekhposyan (2016) propuseram o seguinte teste de racionalidade de flutuação:

$$\max_t \mathcal{W}_{t,m} \quad \text{onde} \quad \mathcal{W}_{t,m} = m \hat{\theta}'_t \hat{V}_\theta^{-1} \hat{\theta}_t \quad \text{para } t = m, m + 1, \dots, P \quad (25)$$

é o teste de Wald em regressões calculadas no tempo t em janelas móveis de tamanho m e é baseado na estimativa de parâmetro $\hat{\theta}_t$, que é estimado sequencialmente na regressão (24), e \hat{V}_θ é um estimador HAC da variância assintótica do parâmetro estimado $\hat{\theta}_t$ na mesma janela rolante.

A hipótese nula é rejeitada se $\max_t \mathcal{W}_{t,m} > \kappa_{\alpha,\ell}$, sendo $\kappa_{\alpha,\ell}$ o valor crítico no nível de significância de $100\alpha\%$, com o número de restrições iguais a ℓ .

A Figura 10 apresenta a trajetória temporal das estatísticas de teste $\mathcal{W}_{t,m}$ com o valor crítico da hipótese nula de racionalidade de previsão para as previsões $t + 1$, $t + 3$, $t + 6$ e $t + 12$ dos modelos com $m = 24$ como referência. O teste de racionalidade em ambientes instáveis rejeita todos os modelos. Como a hipótese nula do teste subjacente é que a previsão é sempre racional, isso não implica irracionalidade na média ou mesmo na maioria dos períodos.

Figura 10 - Teste de Racionalidade da Previsão (Rossi e Seknposyan, 2016) (continua)

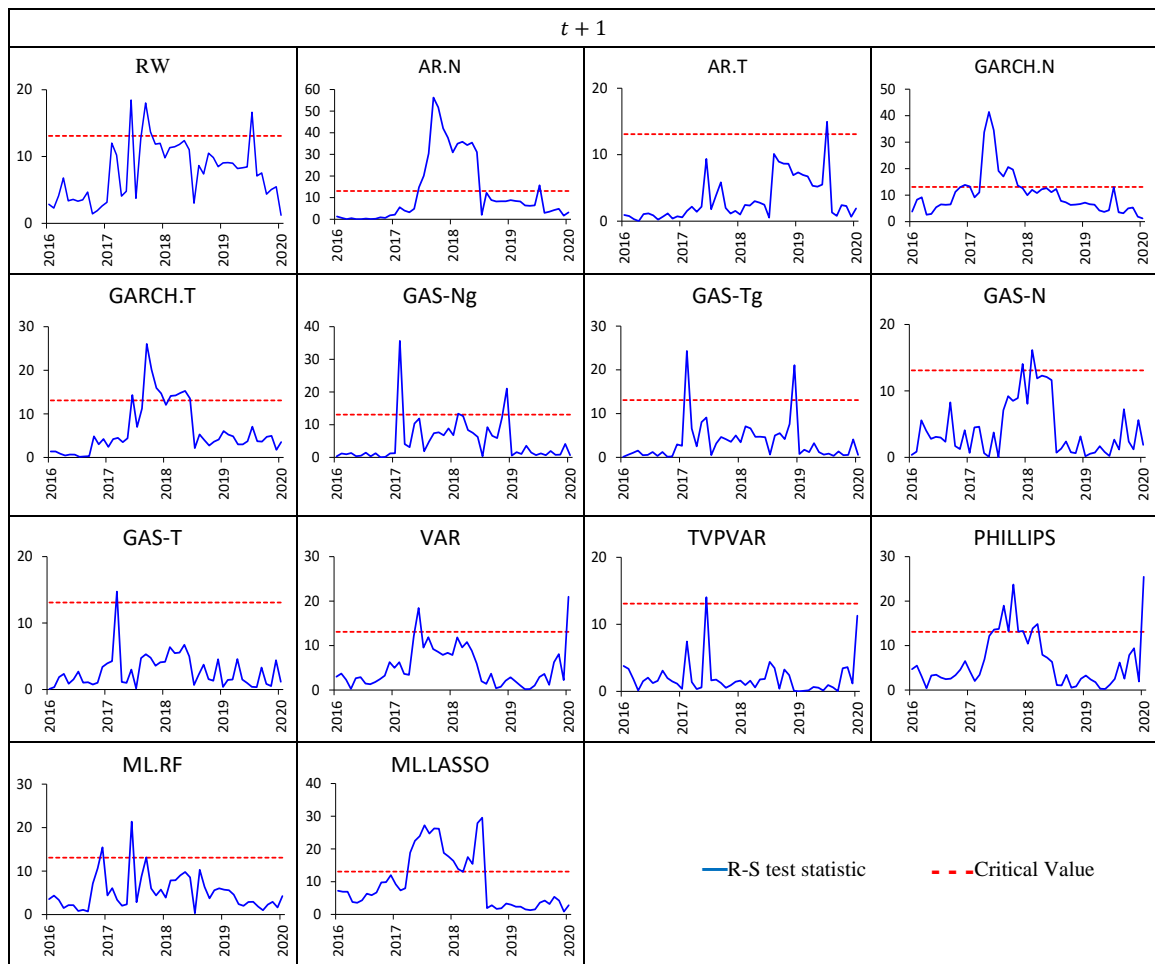


Figura 10 - Teste de Racionalidade da Previsão (Rossi e Seknposyan, 2016) (continuação)

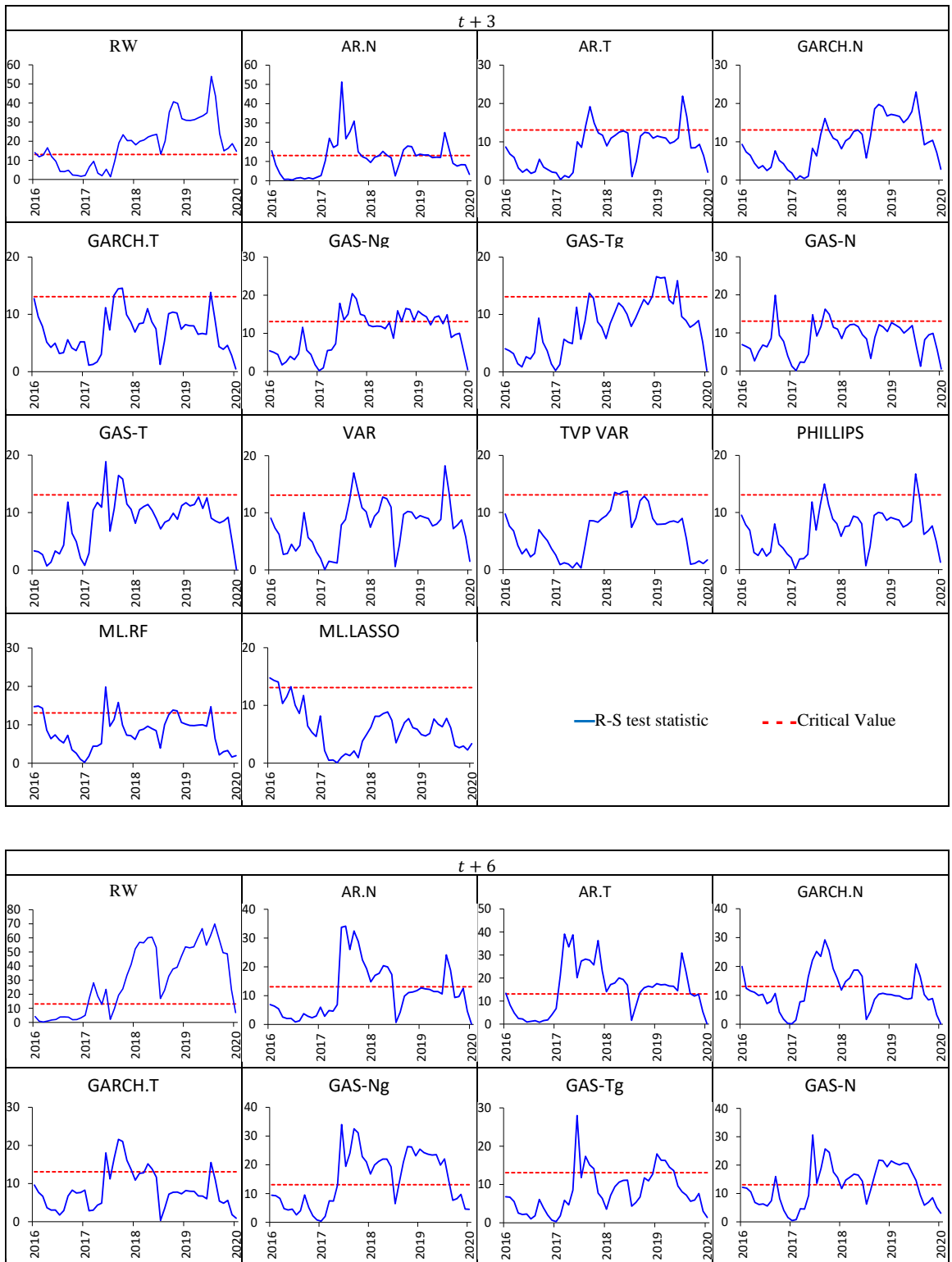
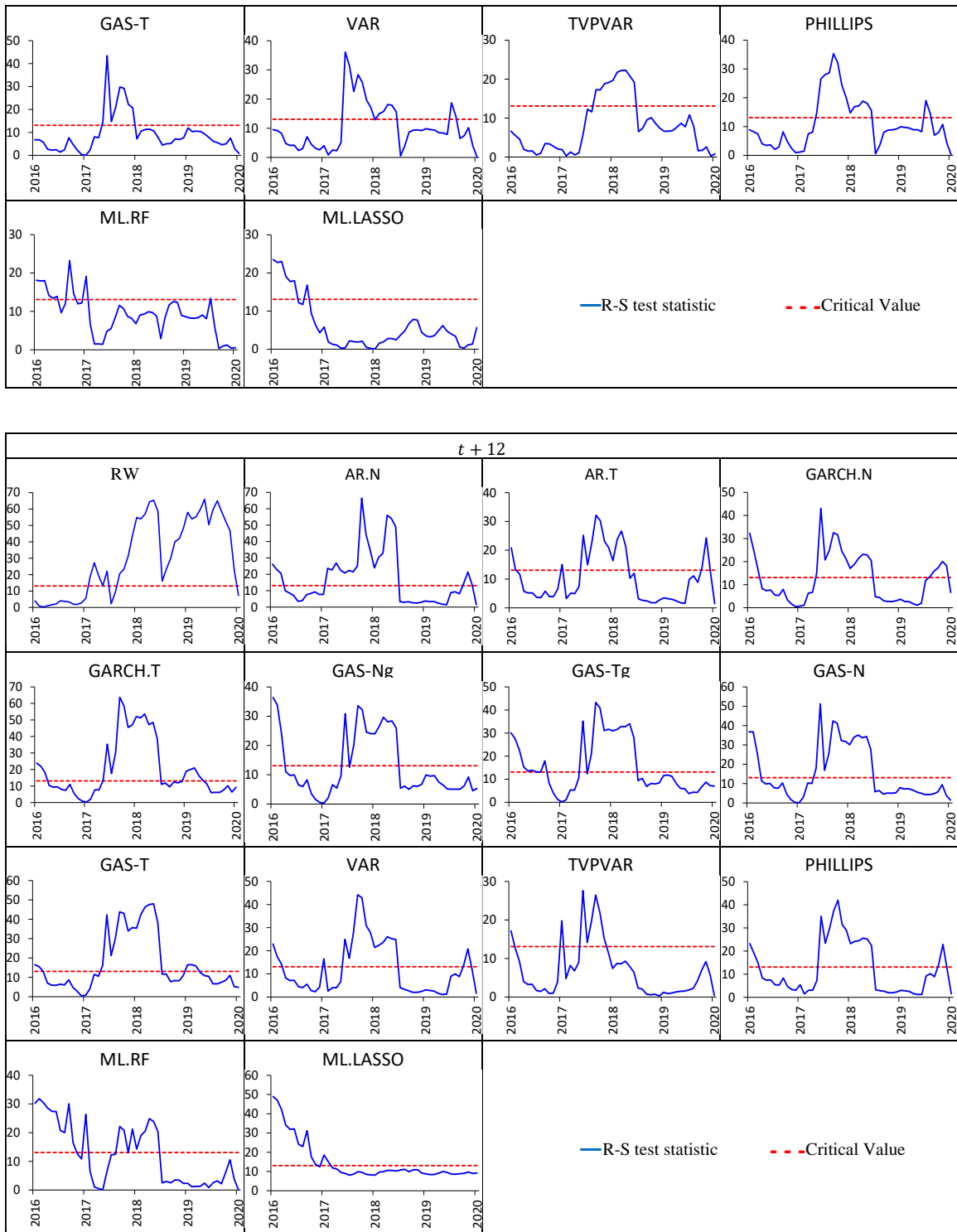


Figura 10 - Teste de Racionalidade da Previsão (Rossi e Seknposyan, 2016) (conclusão)



Fonte: O autor, 2022.

Nota: A linha tracejada apresenta o valor crítico do teste de racionalidade de flutuação ao nível de significância de 5%. A linha contínua relata a estatística de teste para previsões de modelos para cada janela móvel. O tempo no eixo x corresponde ao ponto final da janela.

Uma inspeção visual da série temporal das estatísticas individuais de Wald que fundamentam a estatística de teste indica que a rejeição é motivada principalmente por um forte viés em 2018, que está relacionado principalmente à instabilidade do IPCA causada pela greve dos caminhoneiros.

Mesmo essa evidência preliminar, baseada em testes de racionalidade, dá um suporte bastante forte à nossa hipótese principal; isto é, incerteza e instabilidade de qualquer forma, afeta fortemente as estimativas dos modelos de previsão.

Em geral, os modelos GAS com distribuição Student-t são um pouco mais estáveis e atingiram valores de estatística de teste mais baixas que os modelos competidores. Apesar desses indicativos de superioridade dos modelos GAS com distribuição *t-Student*, a hipótese nula foi rejeitada. Neste caso, será necessário realizar uma comparação direta dos modelos GAS com os modelos competidores através do teste de flutuação de Giacomini e Rossi (2010).

1.7.2.2. Teste de flutuação de Giacomini e Rossi (2010)

Para entender melhor o padrão dos resultados obtidos e verificar em que grau o desempenho preditivo relativo varia ao longo do período fora da amostra, foi aplicado o teste de flutuação desenvolvido por Giacomini e Rossi (2010) e exploradas as evidências de desempenho preditivo local. Os autores argumentam que o desempenho preditivo relativo entre os modelos pode variar substancialmente ao longo do período fora da amostra e calcular a média de tais movimentos para calcular EAM ou EQM pode resultar em perda de informações. Por exemplo, o profissional pode ignorar o fato de que a especificação com o EQM mais alto calculado durante todo o período fora da amostra realmente produz previsões mais precisas ao considerar dados recentes.

Assim como o teste de racionalidade em ambientes instáveis é o valor máximo de estatísticas de teste de racionalidade individual em uma janela móvel, o teste de desempenho de previsão relativa em ambientes instáveis de Giacomini e Rossi (2010) é o valor máximo dos testes tradicionais de desempenho de previsão relativa em uma janela móvel ao longo do tempo. Giacomini e Rossi (2010) propuseram um teste de flutuação que usa a estatística de teste de Diebold e Mariano (1995), mas é calculado em janelas móveis de m anos. Eles derivaram a distribuição assintótica do valor máximo dessa estatística, em todas as janelas possíveis.

O teste de flutuação compara os desempenhos de previsão relativos de modelos concorrentes ao longo do tempo, onde o desempenho é julgado com base em uma função de perda escolhida pelo previsor. Seja $L(\cdot)$ a função de perda escolhida pelo pesquisador e seja $L^{(j)}(\cdot)$ a perda correspondente ao modelo j , $j = 1, 2$. O pesquisador pode usar uma sequência de P diferenças de perda de previsão fora da amostra, $\{\Delta L_{t,h}\}_{t=1}^P$, onde $\Delta L_{t,h} \equiv L_{t,h}^{(1)} - L_{t,h}^{(2)}$, que dependem das realizações da variável π_{t+h} . Por exemplo, para a perda quadrática tradicional associada às medidas de erro quadrático médio de previsão, $L_{t,h}^{(1)} = v_{t+h}^2$, e $\Delta L_{t,h}$ é a diferença entre os erros quadráticos de previsão dos dois modelos concorrentes. Como a função de perda quadrada é a função de perda mais amplamente usada na prática, foi implementada no procedimento descrito abaixo.

Giacomini e Rossi (2010) definiram a perda relativa local para os dois modelos como a sequência de diferenças de perda fora da amostra calculadas em janelas móveis de tamanho m :

$$m^{-1} \sum_{j=t-m+1}^t \Delta L_{j,h} \quad t = m, m+1, \dots, P \quad (26)$$

Testa-se a hipótese nula de mesma capacidade preditiva em cada ponto no tempo, $H_0: E(\Delta L_{t,h}) = 0, \forall t$, onde a alternativa pode ser $E(\Delta L_{t,h}) \neq 0$ (alternativa bilateral) ou $E(\Delta L_{t,h}) > 0$ (alternativa unilateral).

Quando se considera a alternativa bilateral, sua estatística de teste de flutuação é o maior valor sobre a sequência das perdas de erro de previsão relativas (redimensionadas) definidas em (27):

$$\max_t |F_{t,m}^{OOS}|, \quad \text{onde} \quad F_{t,m}^{OOS} = \hat{\sigma}^{-1} m^{-1/2} \sum_{j=t-m+1}^t \Delta L_{j,h} \quad t = m, m+1, \dots, P \quad (27)$$

e $\hat{\sigma}^2$ é um estimador de heterocedasticidade e autocorrelação (HAC) consistente da variância de longo prazo das diferenças de perda (Newey e West, 1987). A hipótese nula é rejeitada contra a hipótese alternativa bilateral $E(\Delta L_{t,h}) \neq 0$ quando $\max_t F_{t,m}^{OOS} > k_{\alpha,\mu}$, onde o valor crítico $k_{\alpha,\mu}$ depende da escolha de μ , que é o tamanho da janela móvel em relação ao número de diferenças de perda fora da amostra.

A superioridade do modelo GAS sob uma distribuição *t-Student* nem sempre é estável em toda a amostra de previsão. Por exemplo, a robustez pode potencialmente atrasar a atualização dos parâmetros na presença de uma quebra estrutural isolada. Neste caso, um modelo Gaussiano adapta-se mais rapidamente ao novo regime, enquanto o modelo *t-Student* inicialmente pondera menos as observações que ocorrem após o intervalo. O objetivo é entender se as melhorias no caso *t-Student* são impulsionadas por episódios particulares ou períodos de amostragem específicos, em vez de serem estáveis ao longo do tempo. A Figura 11 mostra o desenvolvimento das estatísticas de teste individuais subjacentes de modelos GAS em relação aos modelos de referência ao longo do tempo.

A evidência empírica a favor dos modelos GAS contra a hipótese nula parece estar presente em meados de 2016 e início de 2019, quando as estatísticas de teste estão abaixo do valor crítico. Em outras ocasiões, o desempenho da previsão é quase idêntico.

Figura 11 - Teste de flutuação de Giacomini e Rossi (2010) (continua)

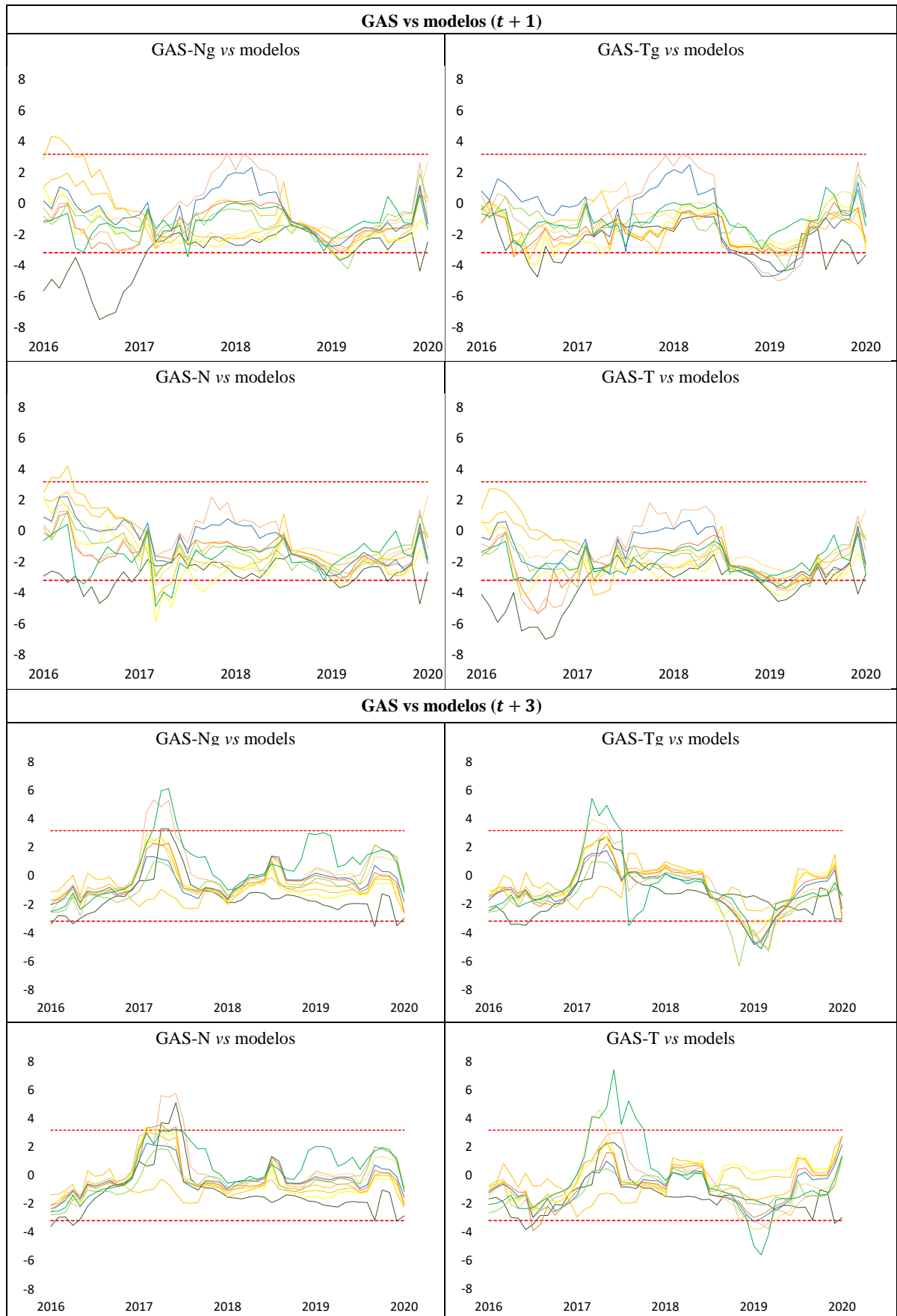
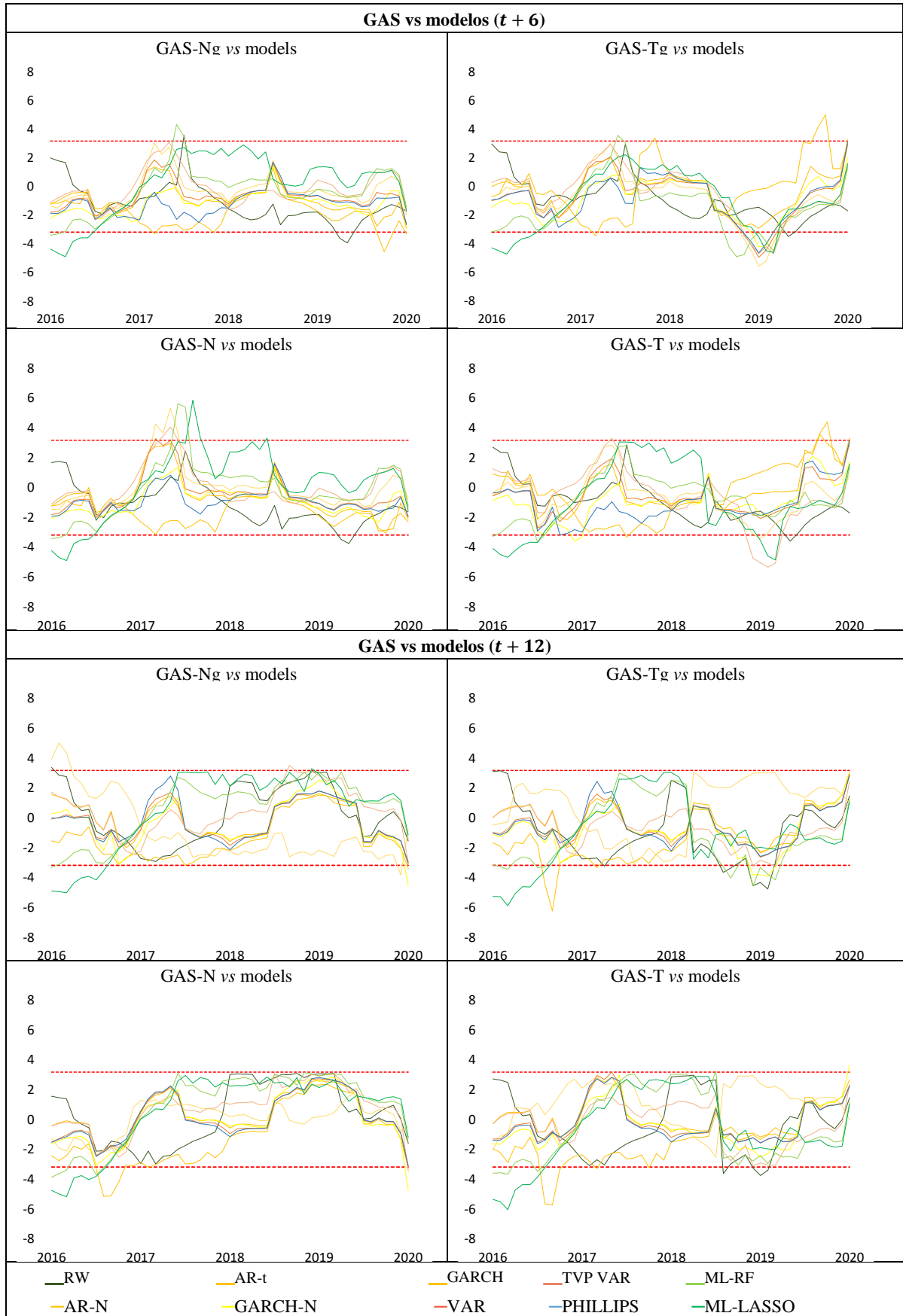


Figura 11 - Teste de flutuação de Giacomini e Rossi (2010) (conclusão)

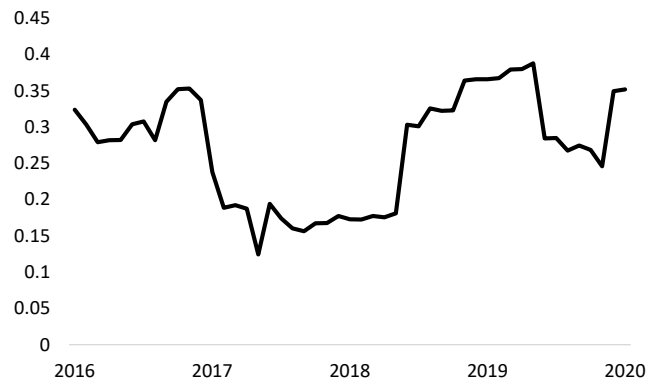


Fonte: O Autor, 2022.

A Figura 12 reporta o desvio padrão da série do IPCA nos últimos 12 meses para o período fora da amostra. Observe que os períodos de volatilidade estão em 2016 e entre meados de 2018 e meados de 2019, indicando maior instabilidade.

A conclusão mais importante do teste é que para períodos de alta instabilidade da série do IPCA, ou seja, quando a previsibilidade era mais difícil, os modelos GAS tiveram melhor desempenho do que os modelos competidores. Assim como em Delle Monache e Petrella (2017) e em de Castro e Aiube (2022), a inclusão da distribuição *t-Student* foi importante para um ganho de performance dos modelos GAS.

Figura 12 - Desvio padrão da série do IPCA



Fonte: O Autor, 2022.

1.7.3. Densidade preditiva

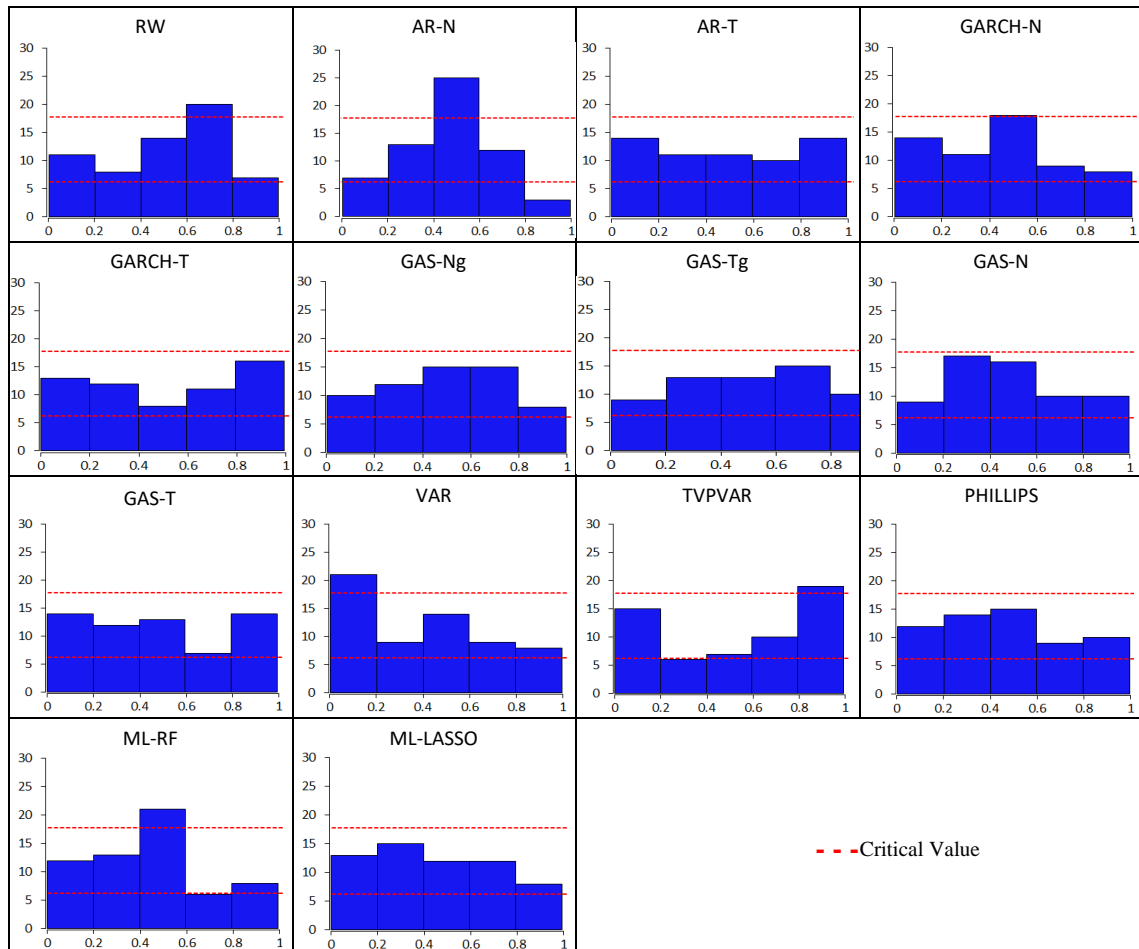
As previsões pontuais estimadas em um esquema de janela móvel, apresentadas na Subseção anterior, fornecem uma boa indicação de qual modelo, em média, é mais preciso. Outra forma de avaliação é através da densidade preditiva, que gera uma métrica de risco medindo a incerteza da previsão. Assim, é possível avaliar qual modelo fornece a melhor caracterização da distribuição de previsão. Isso é relevante para atribuir probabilidades confiáveis a todas as realizações futuras possíveis. Por exemplo, Cogley e Sargent (2015) destacaram a relevância do risco de deflação, cuja previsão requer uma estimativa da densidade geral.

Neste trabalho, foram geradas as densidades preditivas usando o método bootstrapping para resíduos dentro da amostra (ver Garcia et al., 2017). Para cada modelo, em cada janela móvel, selecionamos aleatoriamente 100.000 observações dos resíduos na amostra e as adicionamos à predição pontual, resultando em uma densidade preditiva empírica.

A consistência estatística das distribuições de previsão e das observações é um requisito central para previsões probabilísticas. A calibração ou confiabilidade das previsões é determinada pela capacidade de um modelo de identificar corretamente sua própria incerteza ao fazer previsões. Em um modelo com calibração perfeita, os dados observados em cada ponto de tempo parecem ter vindo da distribuição de probabilidade preditiva naquele momento. Conforme destacado por Diebold et al. (1998), uma densidade preditiva condicional corretamente calibrada produz *probability integral transforms* (PIT) distribuídas

uniformemente. Para o caso de um passo à frente, plota-se na Figura 13 a qualidade da densidade do procedimento PIT de Diebold et al. (1998) para verificar se a hipótese de distribuição condicional é confiável testando se os PITs são uniformes no intervalo (0, 1). O teste rejeita a uniformidade para os modelos RW, AR-N, VAR TVP VAR, ML-RF. Esses modelos superestimam ou subestimam as realizações em alguns intervalos em relação ao que seria esperado se os PITs fossem i.i.d. uniforme.

Figura 13 - Histograma de PITs para previsão um passo à frente



Fonte: O Autor, 2022.

Nota: Histograma dos PITs (normalizados) para todos os modelos de previsões um passo à frente. As linhas vermelhas denotam o intervalo de confiança de 95%.

Seguindo Delle Monache e Petrella (2017), para avaliação da calibração da densidade preditiva de um passo à frente, também foi aplicado o teste não paramétrico de Hong e Li (2005) e o teste de distribuição uniforme de Kolmogorov-Smirnov. Já para horizontes de previsão além de um passo à frente, foi aplicado o teste de Knüppel (KN) (Knüppel, 2015). Este teste verifica a calibração correta das densidades preditivas multi-passo à frente usando momentos brutos (*raw*), ou seja, testa se momentos selecionados dos PITs, por exemplo, são os mesmos que os momentos correspondentes de uma distribuição uniforme. É robusto à presença de correlação serial dos PITs. Sob a hipótese nula, há calibração correta da densidade de previsão vis-à-vis a densidade real. A Tabela 6 relata os resultados da calibração.

O teste KS favorece a especificação correta de todos os modelos, enquanto o teste HL encontra evidências de erro de especificação apenas para os modelos RW e TVP VAR. Para 3

e 6 passos à frente, o teste KN está bem calibrado para todos os modelos. Para 12 passos à frente, muitos modelos rejeitaram a hipótese nula, inclusive todos os modelos GAS.

Também foi medido o desempenho relativo das previsões de densidade para vários modelos. Classificamos as previsões de densidade dos modelos concorrentes com base em três *scores rules* comumente usadas na literatura: logarithmic scores (LogS), *continuous ranked probability scores* (CRPS) and Dawid-Sebastiani score (DSS).

Tabela 6 - Avaliação de previsão de densidade fora da amostra de todos os modelos e todos os horizontes de previsão

Índice de preços ao consumidor brasileiro: fora da amostra																	
	t+1					t+3				t+6				t+12			
	ALS	CRPS	DSS	HL	KS	ALS	CRPS	DSS	KN	ALS	CRPS	DSS	KN	ALS	CRPS	DSS	KN
RW	0.38	1.86	-0.85	3.07*	0.34	0.67	2.59	-0.46	0.19	0.77	2.95	-0.35	0.90	0.56	2.21	-0.74	0.58
AR-N	0.39	1.50	-1.31	1.39	0.76	0.96	1.96	-0.87	0.31	0.87	2.20	-0.86	0.16	0.70	2.17	-1.09	0.03
AR-T	0.49	1.66	-1.31	-0.38	0.50	0.60	2.17	-0.92	0.63	0.58	1.95	-0.90	0.52	0.48	1.95	-1.04	0.44
GARCH-N	0.30	1.70	-1.26	-0.25	0.50	0.42	2.03	-1.08	0.49	0.55	2.14	-1.04	0.42	0.40	1.85	-1.03	0.30
GARCH-T	0.29	1.57	-1.22	1.34	0.34	0.49	1.81	-0.98	0.28	0.49	1.97	-1.05	0.14	0.59	2.19	-1.04	0.07
GAS-Ng	0.87	1.80	-1.38	1.97	0.27	2.58	2.08	-0.63	0.07	1.86	1.93	-0.55	0.08	1.09	1.83	-0.82	0.01
GAS-Tg	0.29	1.51	-1.59	1.06	0.15	1.05	1.99	-0.91	0.09	1.19	2.03	-1.16	0.09	1.11	2.06	-0.92	0.02
GAS-N	1.11	1.63	-1.49	1.59	0.35	2.10	1.91	-0.68	0.08	1.19	1.97	-0.69	0.06	1.08	1.93	-0.96	0.02
GAS-T	0.39	1.47	-1.67	2.20	0.05	1.97	1.96	-1.11	0.06	1.12	1.94	-1.20	0.06	1.08	1.82	-1.05	0.01
VAR	0.35	1.65	-1.41	1.88	0.60	0.90	2.21	-0.85	0.15	1.04	2.30	-0.86	0.10	0.74	1.93	-0.96	0.07
TVP VAR	0.45	1.63	-1.30	2.56*	0.07	1.26	1.95	-0.94	0.29	0.84	1.77	-1.02	0.28	0.76	1.86	-1.21	0.33
PHILLIPS	0.35	1.53	-1.34	1.22	0.48	0.78	2.08	-0.83	0.21	0.96	2.14	-0.87	0.10	0.65	2.29	-0.94	0.04
ML-RF	0.41	1.66	-1.28	0.31	0.40	0.44	1.98	-1.09	0.57	0.56	2.15	-0.95	0.52	0.39	1.98	-1.02	0.41
ML-LASSO	0.43	1.61	-1.38	1.13	0.40	0.76	1.80	-1.10	0.36	0.73	1.97	-0.97	0.32	0.60	2.09	-1.04	0.34

Fonte: O autor, 2022.

Nota: (ALS) - Log score médio; (CRPS \times 10) - Continuous ranked probability score médio; (DSS) - Dawid-Sebastiani score médio. Os modelos indicados em **negrito** são os melhores de acordo com cada previsão de densidade e para cada horizonte de previsão. (HL) - Estatística de teste Hong-Li, (*) rejeita a hipótese nula; (KS) - p-valor do teste de Hong and Li, (KN) - p-valor de teste Knüppel's (2015).

LogS é baseado na evolução ao longo do tempo da probabilidade preditiva logarítmica, ou seja, a probabilidade logarítmica de observar a realização verdadeira da variável. LogS tende a ser sensível à distância entre o centro da densidade de previsão e o resultado realizado. Devido a essa maior sensibilidade a outliers, Gneiting et al. (2005) descobriram que a minimização do LogS pode levar a previsões excessivamente dispersas.

O CRPS é definido em termos de funções de distribuição cumulativa preditiva. Conforme pesquisado em Gneiting et al. (2007), esta regra dos scores (*score rules*) recompensa concentração da densidade preditiva em torno de seu centro, e é menos sensível à distância.

O *score rules* multivariada do DSS (Dawid e Sebastiani, 1999) proposta na literatura dependem apenas dos dois primeiros momentos da distribuição preditiva, ou seja, a média e a covariância.

Suponha que se estime as densidades preditivas $\hat{g}(\cdot)$ para um determinado modelo. Seja π o valor observado da variável IPCA no momento da previsão. LogS, CRPS e DSS são calculados da seguinte forma, respectivamente:

$$\text{LogS}(\hat{g}, \pi) = -\log \hat{g}(\pi) \quad (28)$$

$$\text{CRPS}(\hat{g}, \pi) = \int^{-\infty}^{+\infty} (\hat{g}(z) - l\{\pi \leq z\})^2 dz \quad (29)$$

$$\text{DSS}(\hat{g}, \pi) = -\log \det(\Sigma) - (\pi - \mu)^T \Sigma^{-1} (\pi - \mu) \quad (30)$$

onde $l\{\pi \leq z\}$ denota a função indicadora, que é igual a um se $\pi \leq z$ e zero caso contrário; e μ e Σ denotam a média e a matriz de covariância de $\hat{g}(\cdot)$. Considera-se que as *scores rules* são negativamente orientadas, de modo que se observa um valor menor quando a probabilidade de π ser observado é alta. Para cada modelo, calcula-se os LogS médio, CRPS médio e DSS médio em todas as janelas móveis para cada horizonte de previsão.

Este exercício é semelhante ao realizado por Geweke e Amisano (2010) e Amisano e Geweke (2013), e permite aferir a contribuição de diferentes observações ao longo do tempo. A Tabela 6 mostra os resultados. Os valores em negrito são os valores médios mais baixos de logS (ALS), CRPS e DSS em cada coluna.

No geral, os resultados corroboram os resultados da previsão pontual da Tabela 5. As previsões de um passo à frente permanecem muito mais precisas em comparação com os outros horizontes de tempo.

Na Tabela 6, o resultado é inconclusivo quanto ao melhor modelo em termos de ALS. Já os testes CRPS e DSS indicam dominância do modelo GAS com erros *t-Student* e com heterocedasticidade condicional, exceto para previsões três passos à frente, em que o modelo ML-LASSO tem o menor valor médio de CRPS. Esses resultados são curiosos, pois o CRPS é menos sensível às realizações localizadas nas caudas da densidade preditiva. Com característica, esperava-se um resultado pior em termos de CRPS para os modelos com distribuição *t-Student*, porém o GAS-T foi superior aos demais.

1.8. Discussão

A partir de técnicas de seleção de variáveis preditoras, foram aplicadas várias especificações de modelos com base na abordagem GAS desenvolvida por Creal et al. (2008) e Harvey e Chakravarty (2008) para produzir previsões do índice oficial de preços ao consumidor brasileiro (IPCA), para vários horizontes de tempo diferentes. Também foram estimados outros modelos de referência da literatura como concorrentes, como RW, AR, GARCH, VAR, TVP VAR, curva de Phillips, ML-RF e ML-LASSO, além das projeções do Boletim Focus.

Modelos sob diferentes premissas foram estimados para rastrear os dados a fim de fornecer uma caracterização parcimoniosa da dinâmica da inflação. A presença das inovações *t-Student* se mostrou importante no mecanismo de atualização do TVP. Além disso, os resultados indicam maior estabilidade dos modelos GAS com caudas pesadas. Isso dá mais

robustez às variações de curto prazo da inflação, o que é interessante para uma economia emergente com alta volatilidade, principalmente no que diz respeito às decisões de política econômica.

Os modelos tradicionais de previsão de variáveis macroeconômicas, como os modelos VAR e AR, costumam assumir que as inovações são normalmente distribuídas. Este ponto é particularmente relevante, pois embora a suposição de que os erros sejam gaussianos seja conveniente, ela limita a capacidade do modelo de capturar o comportamento da cauda da distribuição, que está presente em muitas variáveis macroeconômicas. De acordo com Delle Monache e Petrella (2017), permitir caudas pesadas leva a melhorias significativas em termos de ajuste e previsões, e a adoção da distribuição *t-Student* é crucial para obter previsões calibrados. Isso é exatamente o que foi encontrado nesta pesquisa.

Apesar dos modelos GAS se mostraram competitivos, as previsões do Boletim Focus foram significativamente superiores, principalmente no curto prazo. Isso motiva uma investigação mais aprofundada das características que levaram ao Boletim Focus ter um bom desempenho preditivo, sendo referência de previsão de muitas variáveis macroeconômicas do Brasil. Entendendo isso, é possível buscar técnicas de aprimoramento de acurácia preditiva dos modelos deste estudo.

Estudos futuros nessa direção são especialmente importantes para os países emergentes. Os resultados podem ajudar acadêmicos e profissionais a entender os determinantes da inflação e também produzir previsões mais confiáveis. As economias emergentes são mais vulneráveis a períodos de hiperinflação; portanto, as políticas fiscal e monetária devem ser guiadas por informações precisas.

2. ENSAIO II – ESTRATÉGIAS EMPÍRICAS PARA GANHO DE ACURÁCIA DE PREVISÃO

No capítulo anterior, foram estimados os modelos da família GAS e outros 10 modelos competidores amplamente usados na literatura de previsão de inflação. As quatro especificações dos modelos GAS foram: normal e homocedástico, normal e heterocedástico; *t-Student* e homocedástico; *t-Student* e heterocedástico. Este último obteve os melhores resultados em termos de estabilidade, erros de previsão e densidade preditiva.

Comparativamente às previsões do Boletim Focus, todos os modelos estimados neste estudo apresentaram erros de previsão médios maiores. Isso demonstra que o Boletim Focus produz previsões acuradas, corroborando com seu uso como referências do mercado.

Neste capítulo, serão aplicadas técnicas de aprimoramento de acurácia preditiva nas previsões estimadas no ensaio anterior, gerando novas previsões da inflação brasileira e comparadas com as previsões do Boletim Focus.

Primeiramente, foi aplicado o modelo de regressão MIDAS (*Mixed Data Sampling*) de Ghysels et al. (2004) para incorporar informações de alta frequência às previsões estimadas pelos modelos individuais. A introdução de informações semanais do índice de preço ao consumidor semanal (IPC-S), da variação da cotação Tesouro IPCA+ com Juros Semestrais (NTB051535) e variação internacional do preço do petróleo aumentaram a acurácia dos modelos para horizontes de previsão mais curtos um e três meses à frente, porém tiveram pouca relevância para horizontes mais longos.

A próxima etapa foi a incorporação de expectativas inflacionária de especialistas acadêmicos e do mercado via modelo de Black e Litterman (Black e Litterman, 1990) às novas previsões estimadas pelo modelo de regressão MIDAS. O banco de dados utilizado foi o surveys of professional forecasters (SPF) do Banco Central do Brasil. Este banco de dados é formado pelo conjunto de previsões de especialistas que formam o sistema de expectativas do mercado do Banco Central do Brasil. Estes são os dados brutos utilizados pelo Boletim Focus para calcular as projeções divulgadas. Novamente houve ganho de performance nas previsões.

Por fim, com base em um conjunto de confiança, foi aplicada uma ampla gama de métodos de combinação aos componentes de previsão. Após a aplicação de testes de acurácia da previsão pontual e densidades preditiva foi verificado que é possível melhorar a precisão combinando modelos. Segundo Garcia et al. (2017), a combinação de modelos de previsão de inflação com base em conjuntos de confiança pode alcançar desempenhos preditivos superiores aos modelos individuais.

Foi evidenciado que a estratégia de cálculo de projeção do Boletim Focus melhorou o desempenho da previsão. Ou seja, a combinação de modelos a partir de funções simples, mas utilizando apenas os cinco melhores modelos, melhorou o poder preditivo. Ademais, toda estratégia empírica adotada neste trabalho foi capaz de gerar resultados com erros médios de previsão menores que o Boletim Focus.

O Ensaio II se baseia em estratégias e métodos de aprimoramento de acurácia preditiva e está organizado da seguinte forma: a Seção 1 consta a apresentação e aplicação do modelo de regressão de frequências mistas MIDAS; já a Seção 2 apresenta o modelo de Black e Litterman e seus resultados de previsão; a Seção 3 apresenta a estratégia e funções de combinação de modelos de previsão, a Seção 4 conclui a pesquisa.

2.1. Previsão da inflação a partir de dados com frequências mistas

Com o objetivo de melhorar a precisão da previsão da inflação brasileira, serão incorporadas informações de variáveis de alta frequência às previsões de cada modelo da Seção anterior. Para isso, será estimado o modelo de regressão com amostras de variáveis com diferentes frequências dentro de uma estrutura de regressão MIDAS (*Mixed Data Sampling*) desenvolvida por Ghysels et al. (2004). A seguir serão apresentadas as equações do modelo, a forma de selecionar a melhor especificação e a estratégia para obter ganhos de acurácia na previsão.

2.1.1. Regressão MIDAS

Ao modelar relações econômicas, é cada vez mais comum encontrar dados amostrados em diferentes frequências. Entre a gama de métodos de previsão usando dados de frequência mista, a regressão MIDAS é considerada um método de previsão simples e direto. De uma forma mais compacta, a regressão MIDAS pode ser expressa como:

$$\alpha(B)y_t = \beta(L)'x_{t,0} + \varepsilon_t \quad (31)$$

Em que

$$\alpha(z) = 1 - \sum_{j=1}^p \alpha_j z^j \quad (32)$$

$$x_{t,0} := \left(x_{tm_0}^{(0)}, \dots, x_{tm_i}^{(i)}, \dots, x_{tm_l}^{(l)} \right)' \quad (33)$$

$$\beta(z) = \sum_{j=0}^l \beta_j z^j \quad (34)$$

$$\beta_j = \left(\beta_j^{(0)}, \dots, \beta_j^{(i)}, \dots, \beta_j^{(l)} \right)' \quad (35)$$

$$L^j x_{t,0} := x_{t,j} = \left(L^j x_{tm_0}^{(0)}, \dots, L^j x_{tm_i}^{(i)}, \dots, L^j x_{tm_l}^{(l)} \right)' \quad (36)$$

Aqui, $\{y_t, t \in \mathbb{Z}\}$ é um processo univariado observado em baixa frequência com B sendo o operador de defasagem de baixa frequência. $\{x_t^{(i)}, t \in \mathbb{Z}\}$, $i = 0, \dots, k$ é a variável observada em maior frequência, com L sendo o operador de defasagem de maior frequência. l denota uma única ordem dos polinômios de defasagem. Se alguns componentes de $\beta(z)$ forem

de ordem inferior, é fácil definir o restante dos coeficientes do polinômio para zero. O i -ésimo período de alta frequência τ pode ser representado em termos de período de baixa frequência t como $\tau = (t - 1)m_i + j, j = 1, \dots, m_i$.

2.1.2. Restrições funcionais de ponderação

Em alguns casos, o número de defasagens usado para regressão MIDAS pode ser bastante alto. Para contornar esse problema, estima-se uma função paramétrica suficientemente flexível com os chamados hiperparâmetros imposta aos parâmetros originais.

$$\beta_j^i = f_i(j, \varphi_i), \quad j = 0, \dots, l_i, \quad \varphi_i = (\varphi_1^{(i)}, \dots, \varphi_{q_i}^{(i)}), \quad q_i \in \mathbb{N} \quad (37)$$

Usualmente, o número de hiperparâmetros é baixo e facilita a estimativa. É conveniente tomar β_j^i da seguinte forma:

$$\beta_j^i := \frac{f(j; \varphi_i)}{\sum_{k=0}^{l_i} f(k; \varphi_i)} \quad (38)$$

Ghysels et al. (2004) defendem uma função de ponderação parcimoniosa para $f_i(j, \varphi_i) = \omega_i(\varphi)$ para $i = 1, \dots, l_i$, com base nas parametrizações exponencial de Almon lag ou Beta lag. O Almon Lag exponencial de ordem q é especificado como

$$\omega_i(\varphi) = \frac{\exp(\varphi_1 i + \varphi_2 i^2 + \dots + \varphi_q i^q)}{\sum_{i=1}^{l_i} \exp(\varphi_1 i + \varphi_2 i^2 + \dots + \varphi_q i^q)} \quad (39)$$

para algum vetor de parâmetro q -dimensional $\varphi = (\varphi_1, \dots, \varphi_q)'$. Já Beta lag é especificado como

$$\omega_i(\varphi) = \frac{\text{Beta}\left(\frac{i}{l_i}; \varphi_1, \varphi_2, \dots, \varphi_q\right)}{\sum_{i=1}^{l_i} \text{Beta}\left(\frac{i}{l_i}; \varphi_1, \varphi_2, \dots, \varphi_q\right)} \quad (40)$$

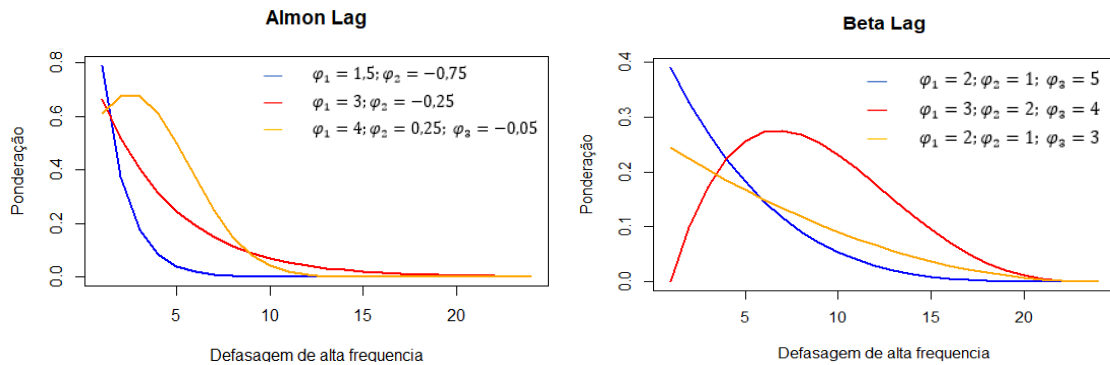
onde $\text{Beta}(\cdot; \varphi_1, \varphi_2, \dots, \varphi_q)$ é a função de densidade de probabilidade de uma distribuição Beta com vetor de parâmetro $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_q)$.

Essas formulações satisfazem a restrição de normalização, ou seja, soma dos pesos igual a um e dado que $\omega_i(\varphi)$ é não negativo, as equações (39) e (40) produzirão valores não negativos.

A Figura 14 apresenta ilustrações de funções de ponderação baseadas em Almon lag e Beta lag de segunda e terceira ordem.

As formas dessas funções de ponderação são muito flexíveis. Os parâmetros da regressão MIDAS podem ser estimados usando mínimos quadrados não lineares ou máxima verossimilhança (ML).

Figura 14 - Funções de ponderação baseadas na exponencial Almon lag e Beta lag



Fonte: O Autor, 2022.

Nota: Os pesos são atribuídos às observações semanais dentro de um mês usando diferentes parâmetros das funções exponencial Almon lag (esquerda) e Beta lag (direita).

Uma alternativa às funções de ponderação Almon e Beta, é a utilização de um parâmetro para cada observação da variável de alta frequência, denominada UMIDAS¹⁰ por Foroni et al. (2011). Nesta abordagem toda informação de X_{tm} é considerada para prever y_t . com a diferença de não restringir que os pesos para as defasagens do regressor estejam ligados por uma função polinomial que pode não ser flexível o suficiente para representar os verdadeiros pesos geradores dos dados aleatórios. O maior prejuízo dessa abordagem está no fato de ela reduzir o número de graus de liberdade da estimação devido ao elevado número de parâmetros, principalmente quando a diferença de frequência entre a variável dependente e os regressores é grande, ou quando o número de defasagens utilizadas para os regressores é elevado.

A restrição exponencial Almon normalizada é uma das restrições mais frequentes escolhidas ao usar o MIDAS. Engle et al. (2013), Colacito et al. (2011) e Andrade et al. (2014) utilizaram a função exponencial Almon como uma das principais restrições utilizadas nas estimativas. Geralmente, leva apenas alguns parâmetros para estimar e é capaz de refletir várias formas funcionais possíveis para os parâmetros. Neste trabalho, além da função de restrição Almon, será utilizada a função Beta.

2.1.3. Seleção de especificações e testes de adequação

Além das considerações usuais sobre as propriedades do termo de erro, há duas questões principais sobre a especificação dos modelos de regressão MIDAS. Primeiro, as restrições funcionais adequadas precisam ser selecionadas, pois sua escolha afetará a precisão do modelo. Em segundo lugar, as ordens de defasagem máxima apropriadas precisam ser escolhidas.

Uma maneira de abordar os dois problemas juntos é usar algum critério de informação (AIC ou BIC) para selecionar o melhor modelo em termos de restrição de parâmetros e ordens

¹⁰ Unrestrict MIDAS

de defasagem. Além disso, é interessante verificar a robustez dos erros padrão à heterocedasticidade e autocorrelação através do teste hAh de Kvedaras e Zemlys (2012). Este teste avalia se uma restrição funcional imposta a parâmetros nas regressões de amostragem de dados mistos é estatisticamente adequada (veja os detalhes em Kvedaras e Zemlys, 2012).

2.1.4. Previsão

Nesta próxima etapa será estimado o modelo de Regressão MIDAS com o principal objetivo de incorporar informações de alta frequência na previsão h -passos à frente da inflação para tentar aumentar sua precisão de cada modelo da Seção anterior.

O conjunto de informações disponíveis no momento t definido por $\mathcal{J}_{t,0} = \{\pi_{t,j}, \pi_{t,j}^{Fq}, x_{t,j}\}_{j=0}^{\infty}$, em que é um vetor de parâmetros dos termos autorregressivos.

$$\begin{aligned}\pi_{t,j} &= (\pi_{t-j}, \dots, \pi_{t-j-p+1})' \\ \pi_{t,j}^{Fq} &= (\pi_{t-j}^{Fq}, \dots, \pi_{t-j-p+1}^{Fq})' \\ X_{t,j} &= (x_{tm_0}^{(0)}, \dots, x_{tm_i}^{(i)}, \dots, x_{tm_k}^{(k)})'\end{aligned}$$

Sendo $\pi_{t,j}$ a inflação observada mensal (baixa frequência), $\pi_{t,j}^{Fq}$, a inflação prevista dentro e fora da amostra mensal (baixa frequência) por cada modelo q da Seção anterior, $X_{t,j}$ é o conjunto de dados em alta frequência (semanal).

O modelo (31) pode ser escrito para o período $t + h$ como pode ser formado apoiando-se em um modelo vinculado a uma expectativa condicional correspondente

$$\pi_{t+h}^{\mathcal{M}} = \alpha'_h \pi_{t,0} + \beta_h(L)' x_{t,0} + \varepsilon_{h,t} \quad \mathbf{E}(\varepsilon_{h,t}, \mathcal{J}_{t,0}) \quad (41)$$

Em que $\pi_{t,0} = (\pi_t, \dots, \pi_{t-p+1})'$ e $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)'$ é um vetor de parâmetros autorregressivos. α_h e $\beta_h(L)$ são os respectivos parâmetros específicos do horizonte h . Observe que foi denotado um sobrescrito " \mathcal{M} " para indicar que a nova previsão de π foi estimada a partir da regressão MIDAS. As expectativas condicionais têm uma forma de representação com certas restrições aos polinômios de defasagem originais dos coeficientes. Assim, no caso geral, as restrições adequadas para cada h terão uma forma diferente.

Reformulando a equação (31) para um modelo autorregressivo com o mês contendo 4 semanas e o ano contendo 48 semanas, tem-se:

$$\pi_{t+h}^{\mathcal{M}} = c + \rho \pi_{t+h}^{Fq} + \sum_{j=0}^{l_i} \omega_i(\varphi) X_{4t-2-j} + \varepsilon_{h,t} \quad (42)$$

Em que π_t^{Fq} são os valores estimados do IPCA de cada modelo q dentro e fora da amostra; \mathbf{X}_{4t} denota o vetor de séries de tempo de alta frequência (semanal); ω_i representa os coeficientes de ponderação e l_i é o número de defasagens de \mathbf{X}_{4t} .

Observe que a contribuição da equação (42) é gerar uma nova previsão estática utilizando como variáveis preditoras, as previsões geradas por cada modelo q e outras variáveis de alta frequência. Ou seja, é um procedimento realizado em duas etapas, pois utiliza uma previsão anterior para gerar uma previsão. É uma alternativa útil para o pesquisador que deseja incluir o efeito de variáveis de alta frequência na previsão gerada por modelos que não permitem a estimação com um conjunto de dados de frequências mistas.

Para cada horizonte de previsão, ou seja, $t + 1$, $t + 3$, $t + 6$ e $t + 12$, será estimado uma regressão MIDAS com um conjunto de especificações potenciais definidas por todas as restrições funcionais e ordens defasadas a serem consideradas. E assim são produzidas combinações de previsão de acordo com um esquema de ponderação de previsão especificado.

A variáveis preditoras de alta frequência (semanal) do conjunto de \mathbf{X} selecionadas são: (i) o IPC-S¹¹, a variação da cotação Tesouro IPCA+ com Juros Semestrais (NTB051535)¹² e variação internacional do preço do petróleo¹³. Em particular, há evidências na literatura de que os preços do petróleo têm conteúdo preditivo relevante para a inflação, ver Clark e Terry (2010) e Stock e Watson (2003), por exemplo.

Observe que os valores de previsão de π^F fora da amostra estão disponíveis, pois são as próprias previsões de cada modelo estimado na Seção 1.7, enquanto os valores de \mathbf{X} estão disponíveis somente nas duas primeiras semanas. Espera-se que o modelo MIDAS tenha mais efetividade para ajudar a prever horizontes de previsão mais curtos como $t + 1$ e $t + 3$.

2.1.5. Resultados

Nesta Seção serão apresentados os resultados dos modelos de previsão MIDAS a partir das funções de restrição mencionadas anteriormente. Foram definidas três especificações a partir de três funções de ponderação de restrição: Irrestrito, Almon Lag e Beta Lag.

Sempre que a adequação empírica não puder ser rejeitada em algum nível apropriado de significância para alguns modelos, podemos ainda contar com critérios de informação para fazer a seleção do(s) melhor(es) candidato(s).

A Tabela 7 apresenta as especificações dos modelos de regressão com determinados números de lags que não rejeitaram a hipótese nula de adequação empírica do teste hAh. Aqui também foi adotado o sobrescrito " \mathcal{M} ", indicando que o modelo de previsão é resultante da regressão MIDAS.

¹¹ O Índice de Preços ao Consumidor - Semanal (IPC-S) mede quadrissemanalmente a variação do custo de vida para famílias com renda entre 1 e 33 salários-mínimos mensais. Fonte: Instituto Brasileiro de Economia - Fundação Getulio Vargas (FGV).

¹² Notas do Tesouro Nacional série B (NTN-B), papéis do governo federal que pagam a correção da inflação do IPCA mais juros. Fonte: Bloomberg

¹³ O barril de petróleo tipo Brent. Fonte: Bloomberg

Todos os p-valores dos testes de adequação das restrições funcionais de parâmetros de todos os modelos para todas as funções de restrição não foram rejeitados ao nível de 5% de significância. Neste caso, Ghysels et al. (2016) sugere que seja selecionada as melhores funções de restrição com base em critérios de informação, neste caso, AIC e BIC. A partir desses critérios, observe que os modelos de regressão com função de restrição polinomial Almog lag exponencial foram selecionados.

Tabela 7 - AIC, BIC e hAh p-value dos modelos de previsão para diferentes funções de restrição

Funções de restrição	Irrestrito			Almon Lag			Beta Lag		
	hAh p-value	AIC	BIC	hAh p-value	AIC	BIC	hAh p-value	AIC	BIC
$RW^{\mathcal{M}}$	0.07	207.7	257.8	0.07	167.7	207.1	0.25	188.9	235.7
$AR - N^{\mathcal{M}}$	0.15	184.4	247.4	0.47	164.4	205.8	0.18	189.4	234.2
$AR - t^{\mathcal{M}}$	0.09	177.4	216.8	0.36	160.3	183.4	0.14	177.1	208.5
$GARCH - N^{\mathcal{M}}$	0.23	192.1	225.3	0.11	166.5	202.8	0.32	178.1	229.2
$GARCH - t^{\mathcal{M}}$	0.32	198.9	211.4	0.30	163.5	188.1	0.41	175.9	206.9
$GAS - Ng^{\mathcal{M}}$	0.25	156.9	200.2	0.34	136.0	174.4	0.60	152.1	190.8
$GAS - tg^{\mathcal{M}}$	0.22	160.6	202.6	0.31	135.2	183.2	0.39	155.5	197.5
$GAS - N^{\mathcal{M}}$	0.22	144.9	166.8	0.21	134.0	158.0	0.40	154.1	169.7
$GAS - t^{\mathcal{M}}$	0.13	140.4	163.6	0.15	130.6	154.6	0.37	150.6	167.1
$VAR^{\mathcal{M}}$	0.09	200.3	199.8	0.24	161.8	186.4	0.47	177.2	200.0
$TVP VAR^{\mathcal{M}}$	0.11	188.6	225.7	0.20	163.5	184.8	0.40	182.3	207.2
$Phillips^{\mathcal{M}}$	0.12	196.6	249.1	0.21	159.6	203.7	0.40	183.9	215.9
$ML - RF^{\mathcal{M}}$	0.15	188.7	198.2	0.40	158.8	184.2	0.46	179.5	204.1
$ML - LASSO^{\mathcal{M}}$	0.08	167.1	186.6	0.31	137.0	175.3	0.85	152.6	187.4

Fonte: O autor, 2022.

Nota: Os valores em negrito representam os modelos com os menores AIC/BIC para cada conjunto de diferentes modelos de defasagem. As especificações escolhidas já representam o melhor número de lags de cada modelo

A seguir serão avaliados os erros de previsão de todos os modelos. A Tabela 8 mostra a raiz do erro quadrático médio (EQM) e o erro absoluto médio (EAM). O modelo estimado neste estudo com o menor erro de previsão para cada horizonte é exibido em negrito.

Como esperado, houve ganho de desempenho, principalmente para previsão de um e três passos à frente. Já para seis e doze passos à frente, o ganho de performance foi mínimo. A incorporação de informações de alta frequência referentes às primeira e segunda semanas de cada mês foram relevantes para melhorar a acurácia de previsões de curto prazo, mas não contribuem muito para previsões de mais longo prazo. Esse resultado corrobora com Gorgi et al. (2019), em que o modelo de regressão MIDAS apresentaram bom desempenho para previsão de um passo à frente da inflação. Os autores também utilizaram o preço de petróleo como variável de alta frequência.

Tabela 8 - Erros de previsão

Horizonte de Previsão	$t + 1$		$t + 3$		$t + 6$		$t + 12$	
	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM
RW ^M	2.97	2.3	3.72	2.91	4.02	3.42	4.66	3.62
AR – N ^M	2.67	1.82	2.96	2.28	3.82	2.85	3.94	2.86
AR – t ^M	2.34	1.64	2.93	2.22	3.42	2.77	3.62	3.39
GARCH – N ^M	2.42	1.79	3.07	2.29	3.57	2.93	3.81	2.95
GARCH – t ^M	2.19	1.62	3.02	2.26	3.43	2.83	3.55	2.74
GAS – Ng ^M	2.11	1.43	3.20	2.51	3.3	2.68	3.65	2.91
GAS – tg ^M	2.25	1.5	2.91	2.25	3.36	2.74	3.32	2.69
GAS – N ^M	2.3	1.64	2.92	2.25	3.32	2.72	3.52	2.79
GAS – t ^M	2.03	1.37	2.88	2.23	3.28	2.54	3.31	2.65
VAR ^M	2.23	1.61	2.97	2.28	3.53	2.88	3.58	2.84
TVP VAR ^M	2.3	1.68	3.08	2.48	3.28	2.76	3.51	2.87
Phillips ^M	2.18	1.57	2.89	2.24	3.51	2.89	3.54	2.89
ML – RF ^M	2.19	1.59	2.91	2.21	3.53	2.69	3.82	2.79
ML – LASSO ^M	2.09	1.54	3.09	2.29	3.75	2.73	3.58	2.68
Focus	1.72	1.34	2.90	2.07	3.06	2.32	3.14	2.44
Top5Focus	1.69	1.26	2.85	1.99	3.05	2.29	3.13	2.43

Fonte: O autor, 2022.

Nota: As previsões Focus e Top5Focus são as originais, não sofrendo alterações da Regressão MIDAS. Foram inseridos na tabela apenas para comparação.

Os modelos GAS com distribuição *t-Student*, assim como na Seção anterior, continuam reportando os menores erros quadráticos médios e erros absolutos médios na grande maioria dos casos. Em especial, o Modelo GAS-t continua com o melhor desempenho dentre todos os modelos. As previsões Focus e Top5Focus permanecem superiores para quase todos os casos, com exceção para EQM de t+3 em que o GAS-t apresentou menor valor que o Focus, já o Top5Focus continua superior para todos os casos.

Na próxima Seção serão incorporadas aos resultados dos modelos de regressão MIDAS, as expectativas inflacionárias de especialistas acadêmicos e do mercado via modelo Black Litterman, gerando assim, novas previsões de inflação. Um fato importante é que para construir as previsões Focus e Top5Focus, o boletim Focus do Banco central do Brasil utiliza o banco de dados formado por essas expectativas. Como um dos objetivos deste trabalho é avaliar se é possível estimar previsões mais precisas que o Boletim Focus, é importante que seja utilizado as mesmas informações disponíveis.

2.2. Incorporação de expectativas via modelo Black e Litterman

Na Seção anterior foram geradas novas previsões da inflação brasileira após a incorporação de informações de alta frequência via modelo de regressão MIDAS. O próximo passo será a incorporação de expectativas inflacionária de especialistas acadêmicos e do mercado via modelo Black e Litterman. Este modelo foi originalmente desenvolvido para

alocação de ativos financeiros, porém para este estudo será adaptado para um contexto de inflação, sendo a variação do IPCA, o retorno do único ativo presente no experimento.

Desde a primeira publicação do artigo Black e Litterman (1990), o seu uso ganhou ampla aplicação prática. Especificamente, o modelo oferece a flexibilidade de incorporar ao retorno de equilíbrio de um ativo ou carteira estimada pelo pesquisador às expectativas do mercado utilizando uma abordagem Bayesiana. Através do uso repetido do teorema de Bayes, a informação *a priori* é atualizada após a incorporação de novas informações a cada instante de tempo, gerando o retorno esperado *a posteriori*.

É possível usar qualquer número de expectativas do mercado, que são declarações sobre as previsões dos retornos esperados de qualquer ativo ou carteiras selecionadas. O modelo produz tanto o conjunto de retornos esperados dos ativos quanto os pesos ótimos do portfólio (He e Litterman, 1999). Os resultados obtidos a partir do modelo Black e Litterman, ao contrário da abordagem tradicional, são bastante intuitivos, estáveis e consistentes em relação as expectativas dos investidores (Ojagverdiyeva, 2011). Diferentes autores sugerem distintas metodologias para tratar as expectativas e construir a matriz de covariância (Meucci, 2006, 2010 and Idzorek, 2007).

2.2.1. Apresentação do modelo

A fórmula dada abaixo é conhecida como equação de Black e Litterman (BL)

$$E[R] = ((\tau\Sigma)^{-1} + P'\Omega^{-1}P)^{-1}((\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q) \quad (43)$$

onde $E[R]$ é o novo vetor ($N \times 1$) de retorno esperado (previsão) do IPCA (*a posteriori*), sendo N o número de ativos presentes na equação, τ é um escalar, Σ é a matriz de covariância ($N \times N$) das previsões de cada modelo q da seção anterior, P é uma matriz ($K \times N$) que identifica os ativos envolvidos nas expectativas, sendo K o número total de expectativas. Entende-se por expectativas, as previsões de cada especialista ou pesquisador que compõem o SPF. Ω é uma matriz de covariância diagonal ($K \times K$) de termos de erro das expectativas representando a incerteza de cada expectativa, Π é o vetor de retorno esperado (previsão) ($N \times 1$) de cada modelo q da seção anterior e Q é o vetor de expectativas (vetor coluna $K \times 1$).

A incerteza das expectativas resulta em um vetor de termo de erro aleatório, desconhecido, independente e normalmente distribuído com média 0 e matriz de covariância Ω . Assim, uma expectativa tem a forma $Q + \varepsilon$

$$Q + \varepsilon = \begin{bmatrix} Q_1 \\ \vdots \\ Q_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_k \end{bmatrix}$$

As variâncias dos termos de erro ω formam Ω , onde Ω é uma matriz de covariância diagonal com zeros em todas as posições fora da diagonal. Os elementos fora da diagonal de Ω são zeros porque o modelo assume que as expectativas são independentes. As variâncias dos

termos de erro ω representam a incerteza das expectativas. Quanto maior a variação do termo de erro, maior a incerteza.

$$\Omega = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \omega_k \end{bmatrix}$$

As expectativas expressas no vetor coluna Q são correspondidas a ativos específicos pela matriz P . Cada expectativa resulta em um vetor de linha ($1 \times N$). Assim, K expectativas resultam em uma matriz ($K \times N$). Já para K expectativas de um único ativo, como o IPCA, resulta em uma matriz ($K \times 1$).

Caso Geral

$$P = \begin{bmatrix} p_{1,1} & \cdots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{k,1} & \cdots & p_{k,n} \end{bmatrix}$$

Um único ativo (IPCA)

$$P = \begin{bmatrix} p_{1,1} \\ \vdots \\ p_{k,1} \end{bmatrix}$$

Uma vez definida a matriz P , a variância de cada expectativa individual pode ser calculada. A variância de uma expectativa individual é $p_k \Sigma p_k'$, onde p_k é um vetor linha ($1 \times N$) da matriz P que corresponde à k -ésima expectativa. A variância respectiva de cada expectativa individual é uma fonte importante de informações sobre a certeza, ou falta dela, do nível de confiança que deve ser colocado em uma expectativa.

O escalar τ e a incerteza das expectativas Ω são os parâmetros mais abstratos e difíceis de especificar do modelo. Quanto maior o nível de confiança nas expectativas do mercado, mais próximas estarão as novas previsões do modelo BL dessas expectativas.

Segundo Satchel (2011), o escalar τ é mais ou menos inversamente proporcional ao peso relativo dado às previsões de cada modelo comparativamente às expectativas do mercado. A orientação na literatura para definir o valor de τ é escassa. Tanto Black e Litterman (1992) quanto Lee (2000) sugerem o escalar τ próximo de zero. Por outro lado, Satchell e Scowcroft (2000) dizem que o valor τ é frequentemente definido como 1. Finalmente, Blamont e Firoozy (2003) sugerem o valor de τ é aproximadamente 1 dividido pelo número de observações.

He e Litterman (2002) calibram a confiança de uma expectativa de modo que a razão de ω/τ seja igual à variância da expectativa $p_k \Sigma p_k'$. Neste caso, a matriz de covariância do termo de erro tem a seguinte forma:

$$\Omega = \begin{bmatrix} p_1 \Sigma p_1' * \tau & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & p_k \Sigma p_k' * \tau \end{bmatrix}$$

2.2.2. Dados

Além dos dados de previsão da inflação brasileira gerados pelo modelo MIDAS na Seção anterior, será utilizado a *Surveys of professional forecasters* (SPF) como conjunto das K opiniões de mercado para previsão de inflação brasileira. SPF são usadas como banco de dados para extrair informações porque essas pessoas têm os incentivos adequados para relatar suas previsões verdadeiras (Keane e Runkle, 1990).

O SPF é um banco de dados do Banco Central do Brasil que recolhe informações de expectativas de variáveis macroeconômicas como inflação, crescimento real do PIB e desemprego em vários horizontes de previsão. Esses dados estão disponíveis na plataforma Bloomberg. Neste trabalho foram utilizadas as previsões profissionais de inflação (IPCA) para um, três, seis e doze meses à frente.

A amostra contempla mais de 1000 observações ao longo do tempo e para diferentes analistas e instituições de janeiro de 2015 até dezembro de 2019, ou seja, o período fora da amostra explorado neste estudo.

2.2.3. Resultados

Esta Seção será destinada a avaliar o ganho de performance das previsões dos modelos estimados na Seção anterior após a incorporação do efeito das expectativas dos especialistas via modelo Black e Litterman.

A matriz de covariância das expectativas dos especialistas foi estimada com base nos dados históricos. A abordagem de janela móvel foi utilizada para construir a matriz para 60 meses de previsão, começando em janeiro de 2015 até dezembro de 2019.

Tabela 9 - Desempenho dos modelos fora da amostra para $t + 1, t + 3, t + 6$ e $t + 12$

Horizonte de Previsão	$t + 1$		$t + 3$		$t + 6$		$t + 12$	
	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM
RW ^{BL}	2.71	1.93	3.52	2.52	3.79	2.70	4.42	3.15
AR – N ^{BL}	2.60	1.86	2.81	2.00	3.46	2.47	3.48	2.48
AR – t ^{BL}	2.16	1.54	2.87	2.05	3.30	2.35	3.14	2.24
GARCH – N ^{BL}	2.32	1.66	2.98	2.12	3.39	2.42	3.36	2.40
GARCH – t ^{BL}	2.15	1.53	2.87	2.05	3.37	2.41	3.27	2.33
GAS – Ng ^{BL}	1.86	1.39	3.03	2.16	3.15	2.25	3.26	2.33
GAS – tg ^{BL}	2.00	1.43	2.73	1.94	3.14	2.24	3.11	2.22
GAS – N ^{BL}	2.16	1.54	2.74	1.95	3.11	2.22	3.21	2.29
GAS – t ^{BL}	1.79	1.36	2.72	1.94	3.03	2.17	3.05	2.18
VAR ^{BL}	2.19	1.56	2.82	2.01	3.39	2.42	3.26	2.33
TVP VAR ^{BL}	2.12	1.51	2.94	2.10	3.23	2.31	3.32	2.37
Phillips ^{BL}	2.07	1.48	2.78	1.98	3.33	2.38	3.06	2.18
ML – RF ^{BL}	2.06	1.47	2.84	2.03	3.29	2.35	3.40	2.43
ML – LASSO ^{BL}	2.02	1.44	3.04	2.17	3.59	2.56	3.38	2.41
Focus	1.72	1.34	2.90	2.07	3.06	2.32	3.14	2.44
Top5Focus	1.69	1.26	2.85	1.99	3.05	2.29	3.13	2.43

Fonte: Autor, 2022.

Além disso, foi selecionado um valor de τ igual a 0.02 para expressar a confiança da distribuição a priori. A partir daqui será denotado um sobrescrito “BL” que indicar que essas

previsões foram geradas pelo modelo Black e Litterman. A Tabela 9 reporta o erro quadrático médio (EQM) e o erro absoluto médio (EAM) dos retornos a posteriori da inflação brasileira após a incorporação das previsões dos especialistas.

Comparativamente à Tabela 8, é possível verificar que as expectativas do mercado sobre o IPCA melhoraram o resultado das previsões. A simulação com as opiniões gerou menores erros tornando as previsões mais precisas. Importante notar que, com exceção das previsões um passo a frente, o experimento com o modelo GAS-t superou o Focus e o Top5Focus, para todos os horizontes de previsão.

2.3. Combinação de modelos de previsão

Como observado por Harvey et al. (1998), embora os modelos possam ter erros de previsão próximos, suas projeções podem diferir significativamente. Os modelos analisados neste estudo têm diferentes variáveis preditoras e premissas, que fazem com que suas propriedades sejam distintas e assim a performance de cada um varie ao longo do tempo.

Sob essa perspectiva, abre a possibilidade de melhorar os resultados usando combinações das previsões geradas por cada modelo. Vale lembrar que as projeções do Boletim Focus são combinações de diversos especialistas que utilizaram modelos distintos.

Em um ambiente em que os modelos individuais estão sujeitos a quebras estruturais e especificação incorreta em vários graus, uma estratégia que reúne informações de muitos modelos normalmente tem um desempenho melhor do que métodos que tentam selecionar o melhor modo de previsão.

Pesquisadores investigaram a combinação ótima de modelos, mas descobriram repetidamente que métodos simples de combinação das previsões produzidas por modelos individuais são benchmarks difíceis de superar e geralmente superam esquemas de ponderação mais sofisticados que dependem da estimativa de pesos teoricamente ótimos (Capistrán et al. 2010; Aiolfi e Timmermann, 2006).

Em um trabalho de revisão da literatura, Timmermann (2006, p. 138-139) elenca quatro grandes razões para se usar métodos de combinação de previsões. A primeira, referida simplesmente como diversificação de portfólio, protege o previsor contra problemas gerados por informação assimétrica. Segunda, previsões individuais são afetadas de maneiras diferentes por mudanças estruturais, dependendo de sua amplitude e duração. Ainda, previsões individuais também estão expostas a erros de especificação, já que o processo gerador dos dados não é conhecido. Finalmente, o autor destaca que os métodos individuais podem ser baseados em diferentes funções de perda.

Samuels e Sekkel (2017) avançaram nessa abordagem e utilizaram o conceito de conjunto de confiança (*Model confidence Set - MCS*) de Hansen et al. (2011) para selecionar os modelos a serem incluídos no conjunto estatisticamente superior de melhores modelos. O número de modelos aparados (excluídos) é determinado por um teste estatístico comparando a precisão de cada modelo.

A ideia de aparar o conjunto de modelos potenciais antes da combinação das previsões foi verificada por Timmermann (2006). Ele considerou três regras simples de corte: combinar

apenas os melhores 75%, 50% ou 25% dos modelos. Ele descobriu que o corte agressivo produzia melhores resultados; em outras palavras, incluir menos modelos na combinação levou a melhores previsões.

Samuels e Sekkel (2017) descobriram que ganhos significativos para o método de corte fixo estão restritos a estratégias que excluem agressivamente 80% a 95% dos modelos. Por outro lado, a regra de corte MCS resultou em melhorias significativas de precisão para uma ampla gama de parâmetros que governam o nível de confiança com o qual o conjunto de melhores modelos é identificado. De Castro e Aiube (2022) corroboram com essa investigação empírica ao combinar modelos de previsão da inflação brasileira.

Dado que as previsões do Boletim Focus são o resultado da simples combinação de várias previsões, seu bom desempenho pode estar associado ao ganho de desempenho ao combinar vários modelos individuais. Além disso, as previsões do Boletim Focus têm ganhos de desempenho se a combinação de previsões for aplicada a apenas um número limitado de modelos. Neste caso, quando a combinação é feita apenas entre as cinco melhores previsões (Top5Focus).

Seguindo Samuels e Sekkel (2017), buscou-se melhorar a precisão combinando esses componentes dos modelos. No entanto, antes de realizar a combinação, os modelos foram selecionados pela abordagem do conjunto de confiança do modelo (MCS), restringindo o número de modelos a serem combinados. Os critérios de seleção e como a combinação é feita são apresentados a seguir.

2.3.1. Conjunto de confiança (MCS)

O procedimento desenvolvido por Hansen et al. (2011) consiste em uma sequência de testes estatísticos que permite construir um conjunto de modelos “superiores”, onde a hipótese nula de igual capacidade preditiva (*equal predictive ability* - EPA) não é rejeitada com um certo nível de confiança (α). Os testes estatísticos da EPA são calculados para uma função de perda arbitrária que satisfaça as condições gerais de estacionariedade fraca, o que essencialmente significa que pode-se testar modelos em vários aspectos, por exemplo, previsões pontuais como em Hansen e Lunde (2005) ou a qualidade do ajuste dentro da amostra como em Hansen et al. (2011).

O MCS permite comparar um grande número de modelos simultaneamente. O teste retorna um conjunto de confiança que inclui o melhor modelo com probabilidade $(1 - \alpha)$. O conjunto se torna mais amplo (com mais modelos) à medida que diminuimos α , enquanto altos valores de α podem resultar em um conjunto com apenas um único modelo.

Dado um conjunto inicial de \hat{M}^0 modelos de dimensão m que englobam todos os modelos da seção anterior, a perda diferencial $d_{ij,t}$ entre os modelos i e j pode ser definida como:

$$d_{ij,t} \equiv L_{i,t} - L_{j,t}, \quad i, j = 1, \dots, m \quad (44)$$

onde $L_{i,t}$ e $L_{j,t}$ as funções de perda. E denota-se

$$d_{i,t} = (m - 1)^{-1} \sum_{j \in M} d_{ij,t}, \quad i = 1, \dots, m \quad (45)$$

como a perda do modelo i em relação a qualquer outro modelo j no instante de tempo t . A hipótese da EPA para um conjunto de M modelos candidatos pode ser formulados de duas maneiras:

$$\begin{aligned} H_{0,M}: \mathbb{E}(d_{ij}) = 0 \text{ para todo } i, j = 1, 2, \dots, m & \quad \text{ou} \quad H_{0,M}: \mathbb{E}(d_{i\cdot}) = 0 \text{ para todo } i = 1, 2, \dots, m \\ H_{A,M}: \mathbb{E}(d_{ij}) \neq 0 \text{ para algum } i, j = 1, 2, \dots, m & \quad H_{A,M}: \mathbb{E}(d_{i\cdot}) \neq 0 \text{ para algum } i = 1, 2, \dots, m \end{aligned}$$

Para testar a hipótese da EPA, a seguinte estatística é construída:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}} \quad t_i = \frac{\bar{d}_{i\cdot}}{\sqrt{\widehat{\text{var}}(\bar{d}_{i\cdot})}} \quad (46)$$

onde $\bar{d}_{i\cdot} = (m - 1)^{-1} \sum_{j \in M} \bar{d}_{ij}$ é a perda amostral do i -ésimo modelo em relação às perdas médias entre os modelos no conjunto M , e $\bar{d}_{ij} = (m - 1)^{-1} \sum_{t=1}^m d_{ij,t}$ mede a perda relativa entre os modelos i th e j th. Observe que $\widehat{\text{var}}(\bar{d}_{i\cdot})$ e $\widehat{\text{var}}(\bar{d}_{ij})$ são as estimativas bootstrap de $\text{var}(\bar{d}_{i\cdot})$ and $\text{var}(\bar{d}_{ij})$, respectivamente. Para ambas as hipóteses da EPA, Hansen et al. (2011) propuseram duas estatísticas para serem usadas como regras de decisão:

$$T_{R,M} = \max_{i,j \in M} |t_{ij}|, \quad T_{\max,M} = \max_{i \in M} t_i,$$

onde t_{ij} e t_i são definidos em (46). Enquanto $T_{R,M}$ usa o diferencial de perda entre os modelos i and j , $T_{\max,M}$ usa o diferencial de perda agregado entre os modelos i e j sobre j . Adota-se $T_{\max,M}$, pois é simples e fácil de calcular. A estatística $T_{R,M}$ cria o conjunto comparando todos os modelos dois a dois, o que torna o procedimento mais intensivo.

O procedimento do conjunto de confiança do modelo (MCS) é um procedimento de teste sequencial, que elimina o pior modelo em cada etapa até que a hipótese de igual capacidade preditiva seja aceita para todos os modelos pertencentes a um conjunto de modelos superiores. A seleção do pior modelo é determinada por uma regra de eliminação consistente com a estatística de teste,

$$e_{R,M} = \arg \max_{i \in M} \left\{ \sup_{j \in M} \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}} \right\} \quad e_{\max,M} = \arg \max_{i \in M} \frac{\bar{d}_{i\cdot}}{\sqrt{\widehat{\text{var}}(\bar{d}_{i\cdot})}} \quad (47)$$

O procedimento também produz p-valores para cada modelo a fim de gerar um ranking.

Neste trabalho, o conjunto de confiança será construído de duas formas: estático e dinâmico. No estático, observa-se toda a janela de 60 meses de previsões geradas pelos modelos para $t + 1$, $t + 3$, $t + 6$ e $t + 12$ e então são gerados os p-valores e o ranqueamento. Já no

modelo dinâmico, a cada horizonte de previsão é formado um novo conjunto de confiança em um esquema de janela móvel. Por exemplo, para as previsões para 3 passos à frente, são formados 20 conjuntos de confiança. Neste caso, o ranqueamento é dinâmico com as posições dos modelos podendo se alternar a cada três passos à frente.

Tabela 10 - Conjunto de confiança

Horizonte de Previsão	$t + 1$		$t + 3$		$t + 6$		$t + 12$	
	Rank	MCS p-valor	Rank	MCS p-valor	Rank	MCS p-valor	Rank	MCS p-valor
RW ^{BL}	eliminado		eliminado		eliminado		eliminado	
AR – N ^{BL}	12	0.95	eliminado		eliminado		eliminado	
AR – t ^{BL}	10	0.95	7	0.97	8	0.97	7	0.92
GARCH – N ^{BL}	eliminado		8	0.97	eliminado		9	0.91
GARCH – t ^{BL}	11	0.95	6	0.98	9	0.97	5	0.97
GAS – Ng ^{BL}	5	0.98	2	0.99	1	0.99	8	0.91
GAS – tg ^{BL}	2	0.99	4	0.99	7	0.95	2	0.99
GAS – N ^{BL}	8	0.95	3	0.99	4	0.98	4	0.97
GAS – t ^{BL}	1	0.99	1	0.99	2	0.99	1	0.99
VAR ^{BL}	6	0.97	5	0.98	10	0.93	eliminado	
TVP VAR ^{BL}	9	0.95	9	0.95	3	0.99	3	0.97
Phillips ^{BL}	4	0.98	eliminado		eliminado		eliminado	
ML – RF ^{BL}	7	0.95	10	0.95	6	0.97	eliminado	
ML – LASSO ^{BL}	3	0.99	11	0.95	5	0.98	6	0.97

Fonte: O Autor, 2022.

Nota: A tabela mostra os p-valores do conjunto de confiança do modelo para todos os horizontes de previsão. Os valores são incluídos no conjunto de confiança de $\alpha = 80\%$ ou 20% . Os p-valores podem ser usados para classificar os modelos. Modelos com números de classificação em negrito são os cinco principais modelos.

Os p-valores e o ranking do MCS são apresentados na Tabela 10. Os valores em negrito representam os modelos mais bem ranqueados no conjunto de confiança com $\alpha = 80\%$.¹⁴

O modelo RW foi eliminado do conjunto de confiança estático de todos os horizontes de previsão. Já os modelos GAS ficaram bem ranqueados e permaneceram no conjunto em todos os horizontes. É importante ressaltar que os modelos GAS com distribuição *t-Student* apresentaram p-valores próximos a 1, em especial o GAS-t que ficou em primeiro lugar no ranking em três dos quatro horizontes de previsão verificados.

O próximo passo é investigar se é possível melhorar o desempenho, combinando os modelos neste estudo. Os resultados do conjunto de confiança estático e dinâmicos servem como base para a realização deste procedimento.

2.3.2. Previsões combinadas

¹⁴ Como o intervalo de tempo fora da amostra tem somente 60 observações, foi necessário impor arbitrariamente um valor de alfa grande para restringir a quantidade de modelos dentro do conjunto de confiança. Caso contrário, poucos modelos seriam excluídos.

Existem inúmeras estratégias de combinação de previsão apresentadas na literatura, muitas não têm sustentação teórica, ficando dependente das especificidades dos dados. Por conta disso, vamos seguir o procedimento de Weiss et al. (2018), que testam uma ampla gama de opções de funções de combinações com o objetivo de melhorar a performance dos modelos apresentados anteriormente.

2.3.2.1. Métodos de Combinação

Desde o artigo de Bates e Granger (1969), foram desenvolvidas inúmeras estratégias de combinação para avançar na literatura teórica e empírica. Segundo Weiss et al. (2018), as melhores formas de combinar diferentes previsões não têm fundamentos teóricos, dependendo muito das especificidades dos dados em questão.

Para fixar notações, denota-se $F_{T \times P}$ como a matriz de previsão com dimensão $T \times P$, onde T é o número de linhas e P é o número de colunas (tem-se P previsões em cada instante de tempo). Denota-se f_i como a previsão obtida utilizando o modelo i , $i \in \{1 \dots P\}$.¹⁵ Alguns métodos de combinação requerem uma ordenação dos componentes de previsão. Quando é este o caso, $f_{(i)}$ denota a i – ésima ordem estatística das previsões de componentes. Finalmente, o peso dado à cada previsão é denotado por w_i , e a previsão combinada por f^c .

A seguir serão apresentados os principais métodos de combinação de previsões encontrados na literatura empírica.

Métodos Simples de Combinação

Agora serão apresentadas algumas maneiras simples em que as previsões podem ser combinadas. Estas são consideradas simples, pois não há necessidade de estimar exatamente o peso que cada previsão deve receber na combinação. Apenas usando alguma medida de *location* (por exemplo, a média) da distribuição cross-sectional das previsões individuais.

Métodos frequentemente utilizados para combinações de previsões:

a) Média Simples (AS)

A forma mais intuitiva é a média simples de todas as previsões. Ao longo dos anos, esta abordagem tem-se estabelecido como um excelente benchmark (por exemplo, Genre et al., 2013). A combinação de previsão é dada de forma simples por

$$f^c = \frac{1}{P} \sum_{i=1}^P f_i$$

¹⁵ Para evitar confusão, omite-se o subscrito adicional t que denota a dimensão temporal da previsão.

Clemen (1989) argumentou que esta ponderação das previsões é frequentemente a melhor estratégia. Stock e Watson (2004) confirmam este resultado e criam um termo chamado de "combinação de previsão puzzle". Em um estudo de simulação realizado por Smith e Wallis (2009), os autores tentaram explicar as razões da média simples dos pesos terem frequentemente um desempenho superior ao das técnicas de combinação de previsão mais sofisticadas, atribuindo este surpreendente achado empírico ao efeito do erro de amostra finita na estimativa dos pesos de combinação. Recentemente, Claeskens et al. (2016) argumentam que o ruído de estimação inferior, quando os pesos são determinados e não estimados, ajuda muito a explicar o *puzzle*. Um estudo geral mais detalhado pode ser encontrado em Graefe et al. (2014).

b) Mediana (MED)

Outro método de combinação bastante simples e atrativo é mediana dos componentes de previsão. A mediana não é sensível à outliers, o que pode ser relevante para algumas aplicações. Palm e Zellner (1992) sugerem que a média simples pode não ser um método de combinação adequado quando algumas das previsões são enviesadas. Isso exige o uso de outras medidas de *location* que sejam robustas à outliers. A mediana é um método de combinação atraente usado em uma ampla gama de estudos empíricos (por exemplo, Armstrong, 1989; McNees, 1992; Hendry e Clements, 2004; Stock e Watson, 2004; Timmermann, 2006). A previsão combinada é dada por:

Para P ímpar:
$$f^c = f_{(\frac{P}{2}+0.5)}$$

Para P par
$$f^c = \frac{1}{2} \left(f_{(\frac{P}{2})} + f_{(\frac{P}{2}+1)} \right)$$

c) Média truncada (TA)

Outra medida robusta à outlier, que é comumente utilizada, é a média truncada (por exemplo, Armstrong, 2001; Stock e Watson, 2004; Jose e Winkler, 2008). Usando um fator de truncagem λ , a previsão combinada é calculada como:

$$f^c = \frac{1}{P(1-2\lambda)} \sum_{i=\lambda P+1}^{(1-\lambda)P} f^{(i)}$$

Para este estudo, será utilizado $\lambda = 0,1$ indicando que foram aparados os 10% superiores e inferiores dos extremos das previsões, excluindo-os do cálculo da previsão combinada. A média truncada é uma interpolação entre a média simples ($\lambda = 0$) e a mediana ($\lambda = 0,5$).

d) Média Winsorizada (WA)

Tal como a média truncada, a média winsorizada é uma estatística robusta que é menos sensível a outliers do que a média simples. Em vez de os remover por completo, como na abordagem de média truncada, limita os outliers a um certo nível, permitindo que tenham um

certo grau de influência. Por esta razão, a medida é por vezes preferida, por exemplo, por Jose e Winkler (2008).

Seja λ o fator de truncagem e $K = \lambda P$. A previsão combinada é então calculada como:

$$f^c = \frac{1}{P} \left[Kf_{(K+1)} + \sum_{i=K+1}^{P-K} Kf_{(P-K)} \right]$$

e) *Variância Mínima – Bates e Granger (1969) (BG)*

Proposto por Bates e Granger (1969), essa abordagem baseia-se na teoria de diversificação de portfólio e usa os elementos da diagonal da matriz de erro quadrático médio da previsão estimada para calcular os pesos da combinação

$$f^c = \sum_{i=1}^P f_i' \times \frac{\hat{\sigma}^{-2}(i)}{\sum_{j=1}^P \hat{\sigma}^{-2}(j)}$$

em que $\hat{\sigma}^2(i)$ é o erro quadrático médio de predição estimado do modelo i .

f) *Newbold e Granger (1974) (NG)*

Com base na investigação anterior de Bates e Granger (1969), a metodologia de Newbold e Granger (1974) também extrai os pesos combinados da matriz de erro quadrático médio de previsão estimada.

Sendo Σ a matriz de erro quadrático médio de previsão de $F_{N \times P}$ e e seja um vetor $P \times 1$ de $(1, 1, \dots, 1)'$. O método de Newbold e Granger (1974) é uma minimização restrita do erro quadrático médio de previsão usando a condição de normalização $e'w = 1$. Isto produz a seguinte combinação previsão:

$$f^c = F_{N \times P} \times \frac{\Sigma^{-1}e}{e'\Sigma^{-1}e}$$

g) *Posto inverso (LnvW)*

Na abordagem de posto inverso, sugerida por Aiolfi e Timmermann (2006), os pesos para as combinações são inversamente proporcionais ao posto dos modelos de previsão individuais:

$$w_i = \frac{Posto_i^{-1}}{\sum_{j=1}^P Posto_j^{-1}}$$

A previsão combinada é obtida por

$$f^c = \sum_{i=1}^P f_i' \times w_i$$

Timmermann (2006) aponta que esse método, assim como Bates e Granger (1969), também ignora correlações entre erros de previsão. No entanto, o método é mais robusto para outliers.

Métodos de combinação baseados em regressão

h) Mínimos Quadrados Ordinários (MQO)

Granger e Ramanathan (1984) mostraram que o método BG é equivalente a um modelo de regressão ajustado por mínimos quadrados em que a constante é suprimida e os pesos são restritos a somar um. Nesta abordagem, as previsões individuais são utilizadas como regressores e a previsão combinada é obtida a partir dos pesos e o intercepto encontrados na regressão:

$$y = \alpha + \sum_{i=1}^P w_i f_i + \varepsilon$$

Utilizando uma parte das previsões para treinar o modelo de regressão, os coeficientes do MQO podem ser estimados através da minimização da soma dos erros quadráticos na equação $w_i = \frac{Posto_i^{-1}}{\sum_{j=1}^P Posto_j^{-1}}$. A previsão combinada é então dada por:

$$f^c = \hat{\alpha} + \sum_{i=1}^P \hat{w}_i f_i$$

i) Mínimos Desvios Absolutos (LAD)

Nowotarski et al. (2014) implementaram a função que minimiza os erros absolutos. Esse método tem menor sensibilidade a outliers pois penaliza os modelos que têm erros grandes para algumas observações.

j) Mínimos Quadrados Restritos (CLS)

O CLS minimiza a soma dos erros quadráticos sob algumas restrições adicionais. Especificamente, restringe-se os coeficientes estimados $\{w_i\}$, permitindo apenas soluções positivas: $w_i \geq 0 \forall i$, e somar um: $\sum_{i=1}^P w_i = 1$. Segundo Weiss et al. (2018) restrições adicionais colocam o CLS subotimamente em comparação com o MQO. Faltam-lhe as boas propriedades assintóticas admitidas pela MQO. No entanto, na prática, verifica-se frequentemente um melhor desempenho, especialmente quando as previsões individuais são altamente correlacionadas.

k) *Regressões Completas de Subconjuntos (CSR)*

O modelo proposto por Elliott et al. (2013) segue a seguinte forma: com os P componentes de previsões que servem como preditores no modelo de regressão, formam-se n modelos de regressão, cada um com um subconjunto exclusivo de preditores. O número total de previsões combinadas geradas por modelos de regressão, n , é dado por

$$n = \sum_{i=1}^P \binom{P}{i} = \sum_{i=1}^P \frac{P!}{i!(P-i)!} = 2^P - 1$$

Essas previsões podem ser, posteriormente, padronizadas para criar uma única combinação de previsões usando pesos que são baseados nos critérios de informação das diferentes regressões individuais. Neste trabalho será utilizado: Akaike (AIC), Bayesian (BIC) e Hannan-Quinn (HQ).

Formalmente, o peso dado a cada previsão com base nos critérios de informação é o seguinte:

$$w_i = \frac{\exp\left(-\frac{1}{2\varrho_i}\right)}{\sum_{j=1}^n \exp\left(-\frac{1}{2\varrho_j}\right)}$$

onde ϱ_i é o critério de informação para a previsão i obtida utilizando uma regressão com uma combinação de previsões específicas. O valor de n é fixado como o número de combinações possíveis, e a previsão combinada é dada por:

$$f^c = \sum_{i=1}^n w_i f_{i_i}$$

Note que este é um método de combinação em duas etapas. O primeiro passo é o cálculo de n previsões combinadas f_i usando o método de Regressões Completas de Subconjuntos com as previsões originais P como preditores; a segunda etapa é a combinação destas previsões cujos pesos seguem determinados critérios de informação.

Métodos de combinação baseados em autovetores

Os métodos de combinação de previsão baseados em autovetores, propostos por Hsiao e Wan (2014), são baseados a ideia de minimizar o erro quadrático médio sujeito a uma condição de normalização.

A condição de normalização mais frequente utilizada para este fim é a de exigir que a combinação pesos some um, ou seja, $\sum_{i=1}^P w_i = 1$ (por exemplo, Newbold e Granger, 1974; Timmermann, 2006). Hsiao e Wan (2014) mostram que esta condição de normalização leva a um mínimo restrito do erro quadrático médio (EQM) de previsão e propõem uma condição de normalização que leve a um mínimo irrestrito do EQM:

$$\sum_{i=1}^P w_i^2 = 1$$

Este mínimo irrestrito do EQM é a base das quatro abordagens baseadas em autovetores apresentadas abaixo.

l) *Autovetor padrão (EIG1)*

A abordagem de autovetor padrão recupera os pesos de combinação da matriz de EQM estimada da seguinte forma: Os autovalores positivos P da matriz EQM são organizados em ordem crescente ($\Phi_1 = \Phi_{min}, \Phi_2, \dots, \Phi_P$), e k_i denota o autovetor correspondente a Φ_i . Seja $d_i = e'k_i$ com e sendo um vetor $P \times 1$ de $(1, 1, \dots, 1)'$. A combinação de pesos w é então escolhida correspondendo ao mínimo de $(\frac{\Phi_1}{d_1^2}, \frac{\Phi_2}{d_2^2}, \dots, \frac{\Phi_P}{d_P^2})$, denotados como k_l , como:

$$w = \frac{1}{d_l} k_l$$

A previsão combinada é então obtida como:

$$f^c = \sum_{i=1}^P f_i' w_i$$

m) *Autovetor com correção de viés (EIG2)*

A abordagem de autovetor com correção de viés baseia-se na ideia de que se um ou mais dos modelos produzirem previsões tendenciosas, a precisão do EIG1 pode ser melhorada eliminando o viés. A abordagem padrão é modificada decompondo os erros de previsão em três partes: viés específico do modelo, fatores comuns omitidos de todos os modelos, bem como uma parte idiossincrática que não está correlacionada entre os modelos.

O procedimento de otimização de EIG2 para obter combinação de pesos coincide com EIG1, exceto pela entrada da matriz de EQM, ou seja, após extrair o viés subtraindo as médias de coluna do EQM:

$$w = \frac{1}{\tilde{d}_l} \tilde{k}_l$$

onde \tilde{d}_i e \tilde{k}_i são definidos analogamente a d_i e k_i na abordagem de autovetor padrão.

A previsão combinada é então obtida por:

$$f^c = \alpha + \sum_{i=1}^P f_i' w_i$$

onde o intercepto α corrige o viés potencial.

Além de EIG1 e EIG2 serão utilizadas outras duas funções de combinação seguindo Hsiao e Wan (2014): autovetor padrão truncado (EIG3) e autovetor com correção de viés truncado (EIG4).

2.3.2.2. Resultados

A partir do conjunto de confiança (MCS) de 1, 3, 6 e 12 passos à frente, serão formados dois grupos:

- a) Todos os modelos que não foram eliminados do MSC apresentados na Tabela 10. Exemplo: Para combinação de 3 passos à frente não serão utilizados os modelos RW, AR-N e Phillips.
- b) Os cinco modelos melhor ranqueados de cada horizonte de previsão. Denota-se “best” para este conjunto. Utilizou-se somente cinco modelos para fazer um balizamento com o mesmo número de previsões do Top5 Focus.

A Tabela 11 mostra todas as funções de combinação que serão aplicadas a cada grupo mencionado acima (para mais detalhes, veja Weiss et al., 2018). O “ID” na primeira coluna se refere ao código de identificação de cada função.

Tabela 11 - Breve descrição das principais funções de combinação de previsão (continua)

ID	Funções de combinação de previsão
<i>Métodos Simples de Combinação</i>	
BG	Variância mínima de Bates e Granger (Bates e Granger, 1969)
InvW	Posto inverso (Aiolfi e Timmermann, 2006)
MED	Mediana (Armstrong, 1989; McNeas, 1992; Hendry e Clements, 2004; Stock e Watson, 2004; Timmermann, 2006)
NG	Método de combinação de Newbold e Granger (Newbold e Granger, 1974)
SA	Média Simples (Genre et al., 2013)
TA	Média truncada (Armstrong, 2001; Stock and Watson, 2004; Jose e Winkler, 2008)
WA	Média Winsorizada (Jose e Winkler, 2008).
<i>Métodos de combinação baseados em regressão</i>	
CLS	Mínimos Quadrados Restritos (Babaie-Kafaki and Roozbeh, 2017)
CSR_aic	Regressão completa do subconjunto por critério de informação: Akaike (aic), Bayesiano (bic) e Hannan Quinn (hq) (Elliott et al., 2013)
CSR_bic	
CSR_hq	
LAD	Mínimos Desvios Absolutos (Nowotarski et al., 2014)
MQO	Mínimos Quadrados Ordinários (Crane and Crotty, 1967, Granger e Ramanathan (1984).
<i>Métodos de combinação baseados em autovetores</i>	
EIG1	Autovetor padrão (Hsiao e Wan, 2014)
EIG2	Autovetor com correção de viés (Hsiao e Wan, 2014)

Tabela 11 - Breve descrição das principais funções de combinação de previsão (conclusão)

EIG3	Autovetor padrão truncado (Hsiao e Wan, 2014)
EIG4	Autovetor com correção de viés truncado (Hsiao e Wan, 2014)
<i>Outras funções de combinação de previsão</i>	
sufixo "dyn"	pesos de combinação dinâmicos

Fonte: O autor, 2022.

Utilizando a versão estática e dinâmica de cada combinação¹⁶ foram produzidas um total de 26 previsões combinadas para o período fora da amostra – janeiro de 2015 até dezembro de 2019. A Tabela 12 apresenta os resultados de EQM e EAM para os todos os horizontes de previsão dos modelos combinados de (a) e (b) comparativamente às previsões Focus.

Tabela 12 - Previsão combinada de erros absolutos médios e erros quadráticos médios

	<i>t</i> + 1		<i>Best t</i> + 1		<i>t</i> + 3		<i>Best t</i> +3		<i>t</i> + 6		<i>Best t</i> + 6		<i>t</i> + 12		<i>Best t</i> + 12	
	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM	EQM	EAM
AS	2.00	1.56	1.93	1.50	2.68	2.09	2.60	2.03	2.92	2.27	2.58	2.01	2.61	2.03	2.57	2.00
BG	1.72	1.34	1.67	1.30	3.03	2.36	2.91	2.27	2.96	2.31	2.62	2.04	3.06	2.38	2.97	2.31
InvW	1.74	1.36	1.67	1.30	2.33	1.82	2.22	1.73	2.85	2.22	2.53	1.97	2.57	2.00	2.49	1.94
MED	2.28	1.78	2.20	1.71	3.06	2.38	3.09	2.41	2.83	2.20	2.57	2.00	2.91	2.27	2.87	2.24
NG	1.69	1.32	1.67	1.30	2.42	1.89	2.30	1.79	3.54	2.76	2.71	2.11	2.86	2.23	2.85	2.22
TA	2.00	1.56	1.92	1.50	2.45	1.91	2.37	1.85	2.77	2.16	2.47	1.92	3.14	2.45	3.09	2.41
BG_dyn	1.89	1.47	1.68	1.31	2.96	2.31	2.75	2.14	3.12	2.43	3.06	2.38	2.68	2.09	2.65	2.06
InvW_dyn	2.15	1.68	2.09	1.63	2.93	2.28	2.82	2.20	2.98	2.32	2.73	2.13	3.04	2.37	2.95	2.30
NG_dyn	1.62	1.26	1.54	1.20	2.38	1.85	2.23	1.74	3.15	2.45	2.79	2.17	2.75	2.14	2.73	2.13
CSR_aic	2.06	1.60	1.86	1.45	2.71	2.11	2.55	1.99	3.30	2.57	2.88	2.24	3.21	2.50	2.63	2.05
CSR_bic	2.10	1.64	1.92	1.50	3.04	2.37	2.87	2.24	3.50	2.73	3.07	2.39	3.54	2.76	2.91	2.27
CSR_hq	2.13	1.66	1.92	1.50	3.06	2.38	2.88	2.24	3.43	2.67	2.95	2.30	3.67	2.86	3.01	2.35
CLS	2.02	1.57	2.00	1.56	3.07	2.39	2.98	2.32	3.06	2.38	2.96	2.31	3.12	2.43	3.02	2.35
LAD	2.27	1.77	1.97	1.53	2.74	2.13	2.58	2.01	3.11	2.42	3.08	2.40	3.45	2.69	3.11	2.42
MQO	2.36	1.84	1.65	1.29	3.31	2.58	3.04	2.37	3.65	2.84	3.37	2.63	3.87	3.02	3.13	2.44
CLS_dyn	2.10	1.64	2.10	1.64	2.83	2.20	2.76	2.15	3.15	2.45	2.95	2.30	3.17	2.47	3.07	2.39
LAD_dyn	2.28	1.78	2.02	1.57	3.20	2.49	2.97	2.31	3.45	2.69	3.33	2.59	2.76	2.15	2.74	2.13
MQO_dyn	2.39	1.86	2.21	1.72	2.74	2.13	2.56	1.99	3.55	2.77	3.27	2.55	3.11	2.42	3.03	2.36
EIG1	2.27	1.77	2.20	1.71	2.65	2.06	2.59	2.02	3.11	2.42	3.60	2.80	3.04	2.37	3.19	2.49
EIG2	1.90	1.48	1.82	1.42	2.91	2.27	2.82	2.20	2.96	2.31	3.50	2.73	2.86	2.23	3.05	2.38
EIG3	1.82	1.42	1.80	1.40	2.89	2.25	2.80	2.18	2.96	2.31	2.75	2.14	2.97	2.31	2.89	2.25
EIG4	2.15	1.68	2.13	1.66	3.04	2.37	2.95	2.30	3.34	2.60	3.12	2.43	3.28	2.56	3.18	2.48
EIG1_dyn	2.16	1.68	2.09	1.63	2.69	2.10	2.56	1.99	3.35	2.61	2.92	2.27	3.10	2.42	3.12	2.43
EIG2_dyn	2.29	1.78	2.20	1.71	2.70	2.10	2.62	2.04	3.06	2.38	2.74	2.13	2.95	2.30	2.88	2.24
EIG3_dyn	1.94	1.51	1.91	1.49	2.70	2.10	2.60	2.03	3.11	2.42	2.86	2.23	2.90	2.26	2.81	2.19
EIG4_dyn	2.10	1.64	2.07	1.61	2.97	2.31	2.84	2.21	3.08	2.40	2.81	2.19	3.25	2.53	3.15	2.45
Focus	1.72	1.34	1.72	1.34	2.90	2.07	2.90	2.07	3.06	2.32	3.06	2.32	3.14	2.44	3.14	2.44
Top5Focus	1.69	1.26	1.69	1.26	2.85	1.99	2.85	1.99	3.05	2.29	3.05	2.29	3.13	2.43	3.13	2.43

Fonte: O Autor, 2022.

¹⁶ As funções AS e TA têm somente a versão estática.

Para facilitar a visualização, Tabela 12 foi hachurada de diferentes cores: amarelo (Métodos Simples de Combinação), verde (Métodos de combinação baseados em regressão) e azul (Métodos de combinação baseados em autovetores). Já os valores em negrito representam os métodos de combinação com os menores valores de EQM e EAM.

Fica claro o potencial de ganho de performance ao combinar modelos independente da técnica utilizada. Quase todos os métodos conseguiram reduzir os erros de previsão. Apesar disso, as técnicas aplicadas tiveram resultados distintos.

Como apontado por Hsiao e Wan (2014), os métodos baseados em regressão tendem a produzir previsões mais precisas quando uma ou algumas das previsões individuais são consideravelmente melhores do que o restante, enquanto os métodos baseados em autovetores têm melhor desempenho quando as previsões individuais estão próximas. Nenhum modelo apresentado tem um desempenho excepcionalmente ruim em comparação aos outros, que são condições que tendem a favorecer as abordagens de autovetores. Apesar disso, os métodos de combinação mais simples apresentaram acurácia preditiva superior que ambos os métodos.

Os métodos de combinação mais simples produziram os EQM e EAM mais baixos para todos os horizontes de previsão, inclusive sendo superiores ao Focus e Top5Focus. Em especial, LnvW foi superior para $t + 3$ e $t + 12$, NG_dyn para $t + 1$ e TA para $t + 6$. Isso corrobora com os resultados encontrados por de Castro e Aiube (2022).

Outro ponto a se observar é que para todos os casos, a combinação dos 5 melhores modelos (*best*) ranqueados no conjunto de confiança (MCS) produziu erros de previsão mais baixos quando comparada à combinação de todos os modelos sobreviventes do conjunto de confiança. Esse resultado é particularmente relevante, pois confirma o melhor desempenho preditivo as previsões divulgadas pelo Top5Focus formada pela combinação dos melhores modelos em comparação ao Focus, formadas pela combinação das expectativas de todos os especialistas que enviaram suas projeções ao Banco Central do Brasil.

A Tabela 13 mostra os p-valores do teste GW, lembrando que a hipótese nula é que os dois modelos tenham a mesma precisão de previsão. O teste foi realizado em pares, onde todos os modelos foram comparados com as previsões do Boletim Focus.

Para p-valores > 0.05 , a hipótese nula de igualdade preditiva não é rejeitada ao nível de significância de 5%. Já para p-valores < 0.05 com os valores em negrito e coloridas na cor vermelha, a hipótese nula é rejeitada em favor ao Boletim Focus. Já para o p-valores < 0.05 com valores em negrito e coloridas na cor azul, a hipótese nula é rejeitada contra o Boletim Focus. Para verificar se a previsão do Boletim Focus foi superior ou inferior à previsão gerada pela determinada função de combinação no caso de rejeição da hipótese nula, foram comparados os erros quadráticos médios.

Apesar de muitas funções de combinação apresentarem erros de previsão menores que o Focus para um passo à frente, o teste não rejeita H_0 para maioria dos casos. A exceção ocorre quando o teste é aplicado a alguns métodos de combinação baseados em regressões. Isso mostra que o poder de previsão pontual de curto prazo do Focus é difícil de ser superado. Já para três, seis e doze passos à frente é possível verificar que algumas funções de combinação baseadas em métodos simples superaram significativamente o Boletim Focus ao rejeitar H_0 .

Tabela 13 - P-valor do teste Giacomini e White em comparação com o Boletim Focus

<i>GW</i>	<i>t + 1</i>	<i>best t + 1</i>	<i>t + 3</i>	<i>best t + 3</i>	<i>t + 6</i>	<i>best t + 6</i>	<i>t + 12</i>	<i>best t + 12</i>
AS	0.474	0.553	0.194	0.1516	0.0882	0.2662	0.0318	0.0428
BG	0.616	0.437	0.460	0.4298	0.0586	0.1467	0.7464	0.0503
InvW	0.319	0.329	0.016	0.0060	0.1424	0.0364	0.0040	0.0061
MED	0.503	0.325	0.424	0.1083	0.1188	0.0217	0.0860	0.0968
NG	0.430	0.236	0.146	0.0190	0.1109	0.0644	0.4629	0.0678
TA	0.359	0.546	0.057	0.0180	0.0382	0.0362	0.0969	0.2616
BG_dyn	0.598	0.368	0.556	0.0684	0.0882	0.2662	0.5085	0.1087
InvW_dyn	0.146	0.116	0.707	0.2560	0.1479	0.1656	0.3039	0.4893
MED_dyn	0.158	0.202	0.165	0.6300	0.0293	0.1864	0.4808	0.1318
NG_dyn	0.442	0.584	0.284	0.032	0.0889	0.2177	0.0918	0.6400
TA_dyn	0.056	0.101	0.456	0.0292	0.1488	0.0802	0.0707	0.6023
CSR_aic	0.145	0.112	0.329	0.6810	0.0882	0.2662	0.1695	0.1510
CSR_bic	0.334	0.180	0.325	0.1492	0.1277	0.1536	0.0994	0.1285
CSR_hq	0.124	0.142	0.449	0.0769	0.1282	0.1539	0.0713	0.6273
CLS	0.002	0.007	0.175	0.2618	0.1277	0.1536	0.0522	0.5980
LAD	0.009	0.006	0.536	0.1489	0.1318	0.1505	0.0951	0.4302
MQO	0.002	0.028	0.077	0.6298	0.1347	0.1685	0.4656	0.1348
CLS_dyn	0.127	0.128	0.431	0.1329	0.1110	0.1468	0.1223	0.6736
LAD_dyn	0.043	0.039	0.190	0.6213	0.1221	0.0556	0.2987	0.0529
MQO_dyn	0.009	0.007	0.754	0.6053	0.1479	0.0656	0.4597	0.1331
EIG1	0.501	0.538	0.789	0.6421	0.1410	0.0893	0.3311	0.0834
EIG2	0.115	0.115	0.452	0.1341	0.1174	0.3309	0.4738	0.5771
EIG3	0.127	0.144	0.098	0.1316	0.0856	0.2466	0.1577	0.1515
EIG4	0.140	0.125	0.122	0.5723	0.1448	0.0695	0.4074	0.2339
EIG1_dyn	0.139	0.265	0.151	0.0505	0.1472	0.0648	0.4501	0.2299
EIG2_dyn	0.141	0.133	0.478	0.4640	0.1156	0.3313	0.1773	0.2410
EIG3_dyn	0.474	0.553	0.094	0.0516	0.1554	0.1716	0.7318	0.6428
EIG4_dyn	0.616	0.437	0.460	0.4298	0.1478	0.0697	0.7464	0.5503

Fonte: O Autor, 2022.

No geral, os resultados da densidade preditiva acompanham os resultados da previsão pontual. A Tabela 14 reporta os resultados das regras de scores das previsões geradas pelos métodos de combinação.

Os valores em negrito representam os melhores métodos de combinação na análise de densidade preditiva. Observe que, novamente, houve dominância de desempenho das funções de combinação simples. Isso pode ser verificado pois menores valores de ALS e CRPS médios. Além disso, nota-se que as combinações geradas a partir do conjunto de confiança (MCS) dos 5 melhores modelos (*best*) apresentaram melhores desempenhos em termos preditivos.

Tabela 14 - Log-scores médios (logS) e CRPS médios ($\times 10$) para todos os modelos e todos os horizontes de previsão

	$t + 1$		<i>best t + 1</i>		$t + 3$		<i>best t + 3</i>		$t + 6$		<i>best t + 6</i>		$t + 12$		<i>best t + 12</i>	
	logS	CRPS	logS	CRPS	logS	CRPS	logS	CRPS	logS	CRPS	logS	CRPS	logS	CRPS	logS	CRPS
SA	0.25	1.62	0.23	1.61	0.56	1.95	0.44	1.83	0.53	1.93	0.48	1.84	0.45	1.89	0.56	1.80
BG	0.26	1.64	0.22	1.62	0.58	1.85	0.55	1.84	0.54	1.98	0.82	1.81	0.36	1.80	0.70	1.77
InvW	0.24	1.67	0.33	1.59	1.11	2.03	0.59	1.88	1.06	2.03	0.77	1.80	0.58	1.94	0.48	1.92
MED	0.29	1.70	0.42	1.60	0.64	1.93	0.77	1.86	0.69	1.94	0.75	1.82	0.73	1.97	0.60	1.94
NG	0.41	1.53	0.50	1.61	2.56	2.09	2.08	2.00	2.77	2.27	1.09	2.05	1.15	2.03	1.34	2.04
TA	0.22	1.58	0.23	1.63	0.54	1.95	0.42	1.82	0.52	1.88	0.47	1.75	0.48	1.91	0.58	1.87
BG_dyn	3.25	1.86	0.39	1.57	1.58	2.23	2.52	2.00	1.10	2.20	1.28	2.03	0.82	1.79	1.58	2.07
InvW_dyn	0.22	1.53	0.29	1.52	0.80	1.88	0.63	1.81	0.70	1.92	0.75	1.78	0.55	1.81	0.45	1.87
MED_dyn	0.30	1.65	0.40	1.65	0.66	1.99	0.78	1.83	0.64	2.04	0.71	1.88	0.73	1.82	0.62	1.84
NG_dyn	0.83	1.64	0.57	1.53	2.74	2.03	2.40	1.97	1.51	2.17	1.14	2.04	0.74	1.84	1.19	1.95
TA_dyn	0.27	1.61	0.30	1.61	0.64	1.89	0.43	1.83	0.51	2.04	0.49	1.79	0.45	1.77	0.54	1.77
CSR_aic	0.54	1.91	0.52	1.56	2.42	2.02	1.68	2.00	2.53	2.18	1.24	1.96	2.46	2.22	1.19	1.96
CSR_bic	0.51	1.96	0.49	1.61	2.40	2.11	1.64	1.99	2.54	2.33	1.16	1.90	2.57	2.40	1.15	2.02
CSR_hq	0.52	1.93	0.50	1.58	2.48	2.04	1.61	1.84	2.61	2.15	1.23	2.00	2.48	2.23	1.20	1.92
CLS	0.27	1.58	0.27	1.57	1.88	2.00	1.81	1.81	1.08	1.97	0.70	1.98	0.90	1.86	1.10	2.02
LAD	0.62	1.95	0.39	1.71	1.39	2.20	2.38	2.06	1.81	2.13	1.70	1.95	1.98	2.18	2.28	2.06
OLS	2.92	2.41	0.52	1.53	2.81	2.33	2.17	2.10	4.12	2.52	1.73	2.12	3.25	2.37	2.40	2.12
CLS_dyn	0.25	1.63	0.27	1.60	1.51	1.91	1.71	1.84	0.72	1.97	0.73	1.95	0.92	2.02	1.11	1.92
LAD_dyn	3.20	1.75	0.39	1.59	1.72	2.12	2.42	2.01	1.17	2.15	1.21	1.94	0.72	1.85	1.52	1.96
OLS_dyn	0.65	1.74	0.58	1.63	3.72	2.17	2.77	1.98	1.75	2.22	1.08	2.11	1.12	1.80	1.38	2.09
EIG1	0.26	1.68	0.23	1.54	0.54	1.98	0.42	1.85	0.53	2.03	0.59	1.78	0.48	1.84	0.51	1.86
EIG2	0.27	1.56	0.24	1.60	0.52	1.93	0.43	1.92	0.58	1.85	0.61	1.79	0.48	1.95	0.47	1.84
EIG3	0.42	1.54	0.32	1.55	1.72	1.94	1.74	1.92	1.23	1.89	0.67	1.93	0.93	2.01	1.09	2.04
EIG4	0.44	1.59	0.33	1.59	1.58	1.90	1.75	1.81	1.22	1.95	0.65	1.82	0.97	1.93	1.11	1.88
EIG1_dyn	0.26	1.66	0.24	1.52	0.55	1.97	0.44	1.82	0.58	1.98	0.62	1.92	0.47	1.86	0.49	1.95
EIG2_dyn	0.25	1.55	0.24	1.52	0.55	1.91	0.44	1.88	0.57	1.92	0.60	1.85	0.49	1.86	0.48	1.94
EIG3_dyn	0.43	1.57	0.33	1.52	1.76	2.01	1.70	1.94	1.35	2.00	0.72	1.87	0.97	2.02	1.11	1.99
EIG4_dyn	0.42	1.59	0.31	1.62	1.65	1.90	1.73	1.98	1.26	2.01	0.70	1.85	0.98	1.94	1.06	2.04

Fonte: O autor, 2009.

Esse resultado corrobora os achados de Armstrong (2001) e Stock e Watson (2004). Segundo esses autores, métodos de combinação mais sofisticados só serão melhores do que métodos mais simples se o previsor tiver fortes evidências de que algumas previsões individuais são muito mais precisas do que outras; caso contrário, é difícil superá-los.

CONCLUSÃO

Após a estimação de previsões geradas pelos modelos GAS e modelos concorrentes, e realizada a comparação com as previsões do Boletim Focus, ficou evidenciado o desempenho superior do Boletim Focus em relação aos demais modelos. Por conta disso, este ensaio buscou aplicar técnicas de aprimoramento de acurácia preditiva nas previsões estimadas no ensaio anterior, gerando novas previsões da inflação brasileira e novas comparações com as previsões do Boletim Focus.

Primeiro, foi aplicado o modelo de regressão MIDAS (*Mixed Data Sampling*) de Ghysels et al. (2004) para incorporar informações de alta frequência às previsões estimadas pelos modelos individuais. Houve ganho de acurácia dos modelos para horizontes de previsão mais curtos, um e três meses à frente, porém pouca relevância para horizontes mais longos. Além disso, também houve ganho de performance ao incorporar expectativas inflacionária de especialistas acadêmicos e de mercado via modelo Black e Litterman às novas previsões estimadas pelo modelo de regressão MIDAS.

Foi evidenciado que a estratégia de cálculo de projeção do Boletim Focus melhorou o desempenho da previsão. Ou seja, a combinação de modelos a partir de funções simples, mas utilizando apenas os cinco melhores modelos, melhorou o poder preditivo. Ademais, toda estratégia empírica adotada neste trabalho foi capaz de gerar resultados com erros médios de previsão menores que o Boletim Focus.

Estudos futuros nessa direção são especialmente importantes para os países emergentes. As economias emergentes são mais vulneráveis a períodos de hiperinflação, portanto as políticas fiscais e monetárias devem ser guiadas por informações precisas. Não obstante, em ambientes inflacionários persistentes como aqueles observados no pós-COVID-19 em países subdesenvolvidos, a precisão das previsões de inflação é importante para direcionar a implementação de políticas públicas. Além disso, destaca-se que os resultados podem ajudar acadêmicos e profissionais a entender os determinantes da inflação.

Esta pesquisa pode ser expandida em várias direções. Diante da grande evolução dos modelos de previsão baseados em *machine learning*, algumas características desta modelagem podem ser incorporadas à dinâmica de evolução dos parâmetros variantes no tempo da abordagem GAS. Isso pouco foi explorado pela literatura até o momento. Ademais, neste estudo a aplicação dos modelos MIDAS e Black e Litterman foi realizada em um procedimento de duas etapas, pois o modelo Black Litterman não permite variáveis com frequências diferentes. Desenvolver uma forma de contornar essa limitação pode ser explorado em pesquisas posteriores.

REFERÊNCIAS

- AIOLFI, M.; CAPISTRÁN, C.; TIMMERMANN, A. Forecast combinations. **CREATES research paper**, n. 2010-21, 2010.
- AIOLFI, M.; TIMMERMANN, A. Persistence in forecasting performance and conditional combination strategies. **Journal of Econometrics**, v. 135, n. 1-2, p. 31-53, 2006.
- AMISANO, G.; GEWEKE, J. Prediction using several macroeconomic models. **Review of Economics and Statistics**, v. 99, n. 5, p. 912-925, 2017.
- ANDRADE, P. et al. The financial content of inflation risks in the euro area. **International Journal of Forecasting**, v. 30, n. 3, p. 648-659, 2014.
- ARESTIS, P.; DE PAULA, L. F.; FERRARI, F. Assessing the economic policies of President Lula da Silva in Brazil: has fear defeated hope. *Political economy of Latin America: recent economic performance*. **Basingstoke: Palgrave Macmillan**, 2007.
- ASCARI, G.; MARROCU, E. **Forecasting inflation: a comparison of linear Phillips curve models and nonlinear time series models**. 2003. CRENoS working paper number 200307.
- ATKESON, A. et al. Are Phillips curves useful for forecasting inflation? **Federal Reserve bank of Minneapolis quarterly review**, v. 25, n. 1, p. 2-11, 2001.
- ARMSTRONG, J. S. Combining forecasts: the end of the beginning or the beginning of the end? **International Journal of Forecasting**, v. 5, n. 4, p. 585-588, 1989.
- ARMSTRONG, Jon Scott. **Principles of forecasting: a handbook for researchers and practitioners**. Boston: Springer, 2001.
- ARRUDA, E. F.; FERREIRA, R. T.; CASTELAR, I. Modelos lineares e não lineares da curva de Phillips para previsão da taxa de inflação no Brasil. **Revista Brasileira de Economia**, v. 65, p. 237-252, 2011.
- BABAIE-KAFAKI, S.; ROOZBEH, M. A revised Cholesky decomposition to combat multicollinearity in multiple regression models. **Journal of Statistical Computation and Simulation**, v. 87, n. 12, p. 2298-2308, 2017.
- BARBOSA-FILHO, N. H. Inflation targeting in Brazil: 1999–2006. *In: Beyond inflation targeting*. Edward Elgar Publishing, v. 22, n. 2, p. 187-200, 2009.
- BARBOSA FILHO, N. O desafio macroeconômico de 2015-2018. **Brazilian Journal of Political Economy**, v. 35, p. 403-425, 2015.
- BATES, J. M.; GRANGER, C. W. J. The combination of forecasts. **Journal of the operational research society**, v. 20, n. 4, p. 451-468, 1969.
- BERGMEIR, C.; HYNDMAN, R. J.; KOO, B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. **Computational Statistics & Data Analysis**, v. 120, p. 70-83, 2018.
- BISAGLIA, L.; GRIGOLETTO, M. A new time-varying model for forecasting long-memory series. **Statistical Methods & Applications**, v. 30, p. 139-155, 2021.
- BLACK, F.; LITTERMAN, R. Asset allocation: combining investor views with market equilibrium. **Goldman Sachs Fixed Income Research**, v. 115, n. 1, p. 7-18, 1990.
- BLACK, F.; LITTERMAN, R. Global portfolio optimization. **Financial analysts journal**, v. 48, n. 5, p. 28-43, 1992.
- BLAMONT, D.; FIROOZY, N. Asset allocation model. **Global Markets Research: Fixed Income Research**, Deutsche Bank, 2003.
- BLASQUES, F.; GORGI, P.; KOOPMAN, S. J. **Accelerating GARCH and Score-Driven Models: optimality, estimation and forecasting**. 2017. Tinbergen Institute Discussion Paper, No. 17-059/III.

- BLASQUES, F.; GORGI, P.; KOOPMAN, S. J. Accelerating score-driven time series models. **Journal of Econometrics**, v. 212, n. 2, p. 359-376, 2019.
- BLAZSEK, S. et al. Forecasting rate of return after extreme values when using AR-t-GARCH and QAR-Beta-t-EGARCH. **Finance Research Letters**, v. 24, p. 193-198, 2018.
- BLAZSEK, S.; CHAVEZ, H.; MENDEZ, C. Model stability and forecast performance of Beta-t-EGARCH. **Applied Economics Letters**, v. 23, n. 17, p. 1219-1223, 2016.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of econometrics**, v. 31, n. 3, p. 307-327, 1986.
- BOX, G. E. P.; JENKINS, G. M. Time series analysis: forecasting and control. **Calif: Holden-Day**. San Francisco, 1976.
- BREIMAN, L. Bagging predictors. **Machine learning**, v. 24, p. 123-140, 1996.
- BREIMAN, L. Random forests. **Machine learning**, v. 45, p. 5-32, 2001.
- CARRARA, A. F.; CORREA, A. L. O regime de metas de inflação no Brasil: uma análise empírica do IPCA. **Revista de Economia Contemporânea**, v. 16, p. 441-462, 2012.
- CASAGRANDE, M. H. Comparação de métodos de estimação para problemas com colinearidade e/ou alta dimensionalidade ($p > n$). 2016 Dissertação (Mestrado em Economia) - Universidade Federal de São Carlos, São Carlos, 2016.
- CLAESKENS, G. et al. The forecast combination puzzle: a simple theoretical explanation. **International Journal of Forecasting**, v. 32, n. 3, p. 754-762, 2016.
- CLARK, T. E.; TERRY, S. J. Time variation in the inflation passthrough of energy prices. **Journal of Money, credit and Banking**, v. 42, n. 7, p. 1419-1433, 2010.
- CLEMEN, R. T. Combining forecasts: a review and annotated bibliography. **International Journal of Forecasting**, v. 5, n. 4, p. 559-583, 1989.
- COGLEY, T.; SARGENT, T. J. The conquest of US inflation: learning and robustness to model uncertainty. **Review of Economic Dynamics**, v. 8, n. 2, p. 528-563, 2005.
- COLACITO, R.; ENGLE, R. F.; GHYSELS, E. A component model for dynamic correlations. **Journal of Econometrics**, v. 164, n. 1, p. 45-59, 2011.
- COSTA FILHO, A. E. Inflação e incerteza inflacionária no Brasil. **Economia Aplicada**, v. 20, n. 4, p. 355-381, 2016.
- CRANE, D. B.; CROTTY, J. R. A two-stage forecasting model: exponential smoothing and multiple regression. **Management Science**, v. 13, n. 8, p. B-501, 1967.
- CREAL, D.; KOOPMAN, S. J.; LUCAS, A. **A general framework for observation driven time-varying parameter models**. 2008. Tinbergen Institute Discussion paper (No. 08-108/4).
- CREAL, D.; KOOPMAN, S. J.; LUCAS, A. Generalized autoregressive score models with applications. **Journal of Applied Econometrics**, v. 28, n. 5, p. 777-795, 2013.
- DE CASTRO, C. H. D. C.; AIUBE, F. A. L. Forecasting inflation time series using score-driven dynamic models and combination methods: the case of Brazil. **Journal of Forecasting**, v. 42, n. 2, p. 369-401, 2023.
- DELLE MONACHE, D.; PETRELLA, I. Adaptive models and heavy tails with an application to inflation forecasting. **International Journal of Forecasting**, v. 33, n. 2, p. 482-501, 2017.
- DOS SANTOS, D. R.; RIBEIRO, T. C.; SANFINS, M. A. Estimating the Level of the Brazilian Yield Curve Using the Time-Varying Coefficient Model GAS (2, 2) with Gamma Distribution. **International Journal of Economics and Finance**, v. 11, n. 9, p. 1-1, 2019.
- DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**, v. 74, n. 366a, p. 427-431, 1979.

- DIEBOLD, F. X.; GUNTHER, T. A.; TAY, S. A. Evaluating density forecasts with applications to financial risk management. **International Economic Review**, v. 39, n. 4, p. 863-883, 1998.
- DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. **Journal of Business & Economic Statistics**, v. 13, p. 253-265, 1995.
- DWECK, E.; TEIXEIRA, R. A. **A política fiscal do governo Dilma e a crise econômica**. Texto para discussão, v. 1, n. 303, p. 01-42, 2017.
- ELLIOTT, G.; GARGANO, A.; TIMMERMANN, A. Complete subset regressions. **Journal of Econometrics**, v. 177, n. 2, p. 357-373, 2013.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, v. 50, n. 4, p. 987-1007, 1982.
- ENGLE, R. F.; GHYSELS, E.; SOHN, B. Stock market volatility and macroeconomic fundamentals. **Review of Economics and Statistics**, v. 95, n. 3, p. 776-797, 2013.
- FAUST, J.; WRIGHT, J. H. Forecasting inflation. *In: Handbook of Economic Forecasting*. Elsevier, v. 2, p. 2-56, 2013.
- FERNANDES, M.; MEDEIROS, M. C.; SCHARTH, M. Modeling and predicting the CBOE market volatility index. **Journal of Banking and Finance**, v. 40, n. 1, p. 1-10, 2014.
- FIGUEIREDO, E.; MARQUES, A. Inflação inercial como um processo de longa memória: análise a partir de um modelo ARFIMA-FIGARCH. **Estudos Econômicos**, v. 39, p. 437-458, 2009.
- FORONI, C.; MARCELLINO, M.; SCHUMACHER, C. **U-midas**: midas regressions with unrestricted lag polynomials. 2011. Working Paper Deutsche Bundesbank, Research Centre.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. A note on the group lasso and a sparse group lasso. **arXiv preprint arXiv:1001.0736**, 2010.
- GAGLIANONE, W. P.; ISSLER, J. V.; MATOS, S. M. Applying a microfounded-forecasting approach to predict Brazilian inflation. **Empirical Economics**, v. 53, n. 1, p. 137-163, 2017.
- GARCIA, M. G.; MEDEIROS, M. C.; VASCONCELOS, G. F. Real-time inflation forecasting with high-dimensional models: the case of Brazil. **International Journal of Forecasting**, v. 33, n. 3, p. 679-693, 2017.
- GENRE, V.; KENNY, G.; MEYLER, A.; TIMMERMANN, A. Combining expert forecasts: can anything beat the simple average? **International Journal of Forecasting**, v. 29, n. 1, p. 108-121, 2013.
- GEWEKE, J.; AMISANO, G. Comparing and evaluating Bayesian predictive distributions of asset returns. **International Journal of Forecasting**, v. 26, n. 2, p. 216-230, 2010.
- Ghysels, E., Santa-Clara, P., & Valkanov, R. **The MIDAS Touch**: mixed Data Sampling Regression Models. 2004. Working Paper Dept. of Economics UCLA: Finance.
- GHYSELS, E.; KVEDARAS, V.; ZEMLYS, V. Mixed frequency data sampling regression models: the R package midasr. **Journal of statistical software**, v. 72, p. 1-35, 2016.
- GIACOMINI, R.; ROSSI, B. Forecast comparisons in unstable environments. **Journal of Applied Econometrics**, v. 25, n. 4, p. 595-620, 2010.
- GIACOMINI, R.; WHITE, H. Tests of conditional predictive ability. **Econometrica**, v. 74, n. 6, p. 1545-1578, 2006.
- GNEITING, T.; BALABDAOUI, F.; RAFTERY, A. E. Probabilistic Forecasts, Calibration and Sharpness. **Journal of the Royal Statistical Society B**, v. 69, n. 2, p. 243-268, 2007.
- GNEITING, T.; RAFTERY, A. E. Weather forecasting with ensemble methods. **Science**, v. 310, n. 5746, p. 248-249, 2005.

- GORGI, P.; KOOPMAN, S. J.; LI, M. Forecasting economic time series using score-driven dynamic models with mixed-data sampling. **International Journal of Forecasting**, v. 35, n. 4, p. 1735-1747, 2019.
- GRAEFE, A.; ARMSTRONG, J. S.; JONES JR, R. J.; CUZÁN, A. G. Combining forecasts: an application to elections. **International Journal of Forecasting**, v. 30, n. 1, p. 43-54, 2014.
- GRANGER, C. W.; RAMANATHAN, R. Improved methods of combining forecasts. **Journal of forecasting**, v. 3, n. 2, p. 197-204, 1984.
- HANSEN, P. R.; LUNDE, A.; NASON, J. M. The model confidence set. **Econometrica**, v. 79, n. 2, p. 453-497, 2011.
- HANSEN, P. R.; LUNDE, A. A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? **Journal of applied econometrics**, v. 20, n. 7, p. 873-889, 2005.
- HASTIE, T.; TIBSHIRAMI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: data Mining, Inference and Prediction**. New York: Springer, 2009.
- HASTIE, T.; TIBSHIRANI, R.; WAINWRIGHT, M. **Statistical Learning with Sparsity: The Lasso and Generalizations**. Boca Raton: CRC Press, 2015.
- HASTIE, T.; TIBSHIRANI, R.; TIBSHIRANI, R. J. **Extended comparisons of best subset selection, forward stepwise selection, and the lasso**, arXiv preprint arXiv:1707.08692, 2017.
- HARVEY, A. **Dynamic models for volatility and heavy tails: with applications to financial and economic time series**. New York: Cambridge University Press, 2013.
- HARVEY, A.; LUATI, A. Filtering with Heavy Tails. **Journal of the American Statistical Association**, v. 109, n. 507, p. 1112-1122, 2014.
- HARVEY, A.; CHAKRAVARTY, T. **Beta-t-(E)GARCH**. 2008. Discussion Paper University of Cambridge CWPE 08340.
- HARVEY, D. I.; LEYBOURNE, S. J.; NEWBOLD, P. Tests for forecast encompassing. **Journal of Business & Economic Statistics**, v. 16, n. 2, p. 254-259, 1998.
- HENDRY, D. F.; CLEMENTS, M. P. Pooling of forecasts. **The Econometrics Journal**, v. 7, n. 1, p. 1-31, 2004.
- HE, G.; LITTERMAN, R. **The intuition behind Black-Litterman model portfolios**. 1999. Working paper Goldman Sachs Quantitative Resources Group.
- HONG, Y.; LI, H. Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. **Review of Financial Studies**, v. 18, n. 1, p. 37-84, 2005.
- HSIAO, C.; WAN, S. K. Is there an optimal forecast combination? **Journal of Econometrics**, v. 178, p. 294-309, 2014.
- IDZOREK, T. A step-by-step guide to the Black-Litterman model: incorporating user-specified confidence levels. *In: Forecasting expected returns in the financial markets*. Academic Press, 2007. p. 17-38.
- IVERSEN, Jens et al. Real-time forecasting for monetary policy analysis: the case of Sveriges Riksbank. **Riksbank Research Paper Series**, n. 142, 2016.
- JOSE, V. R. R.; WINKLER, R. L. Simple robust averages of forecasts: some empirical results. **International journal of forecasting**, v. 24, n. 1, p. 163-169, 2008.
- KAPETANIOS, G.; LABHARD, V.; PRICE, S. Forecasting using bayesian and information-theoretic model averaging: an application to UK inflation. **Journal of Business & Economic Statistics**, v. 26, n. 1, p. 33-41, 2008.
- KASCHA, C.; RAVAZZOLO, F. Combining inflation density forecasts. **Journal of forecasting**, v. 29, n. 1-2, p. 231-250, 2010.

- KNÜPPEL, M. Evaluating the calibration of multi-step-ahead density forecasts using raw moments. **Journal of Business & Economic Statistics**, v. 33, n. 2, p. 270-281, 2015.
- KOOPMAN, S. J.; LUCAS, A.; SCHARTH, M. Predicting Time-Varying Parameters with Parameter-Driven and Observation-Driven Models. **The Review of Economics and Statistics**, v. 98, n. 1, p. 97-110, 2016.
- KOHLSCHEEN, E. Uma nota sobre erros de previsão da inflação de curto-prazo. **Revista Brasileira de Economia**, v. 66, p. 289–297, 2012.
- KWIATKOWSKI, D.; PHILLIPS, P. C.; SCHMIDT, P.; SHIN, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? **Journal of Econometrics**, v. 54, n. 1-3, p. 159-178, 1992.
- KVEDARAS, V.; ZEMLYS, V. Testing the functional constraints on parameters in regressions with variables of different frequency. **Economics Letters**, v. 116, n. 2, p. 250-254, 2012.
- LEE, W. **Advanced Theory and Methodology of Tactical Asset Allocation**. New York, NY: John Wiley & Sons, 2000.
- LUCAS, A.; ZHANG, X. Score-driven exponentially weighted moving averages and Value-at-Risk forecasting. **International Journal of Forecasting**, v. 32, n. 2, p. 293-302, 2016.
- MCNEES, S. K. The uses and abuses of ‘consensus’ forecasts. **Journal of Forecasting**, v. 11, n. 8, p. 703-710, 1992.
- MEDEIROS, M.; VASCONCELOS, G.; FREITAS, E. Forecasting Brazilian inflation with high-dimensional models. **Brazilian Review of Econometrics**, v. 36, p. 68–100, 2016.
- MEUCCI, A. **Beyond Black-Litterman in practice: a five-step recipe to input views on non-normal markets**. 2006. Working Paper Series.
- MEUCCI, A. The Black-Litterman approach: original model and extensions. **Shorter version in, The Encyclopedia of Quantitative Finance, Wiley, 2010**.
- MINCER, J. A.; ZARNOVITZ, V. The evaluation of economic forecasts. *In*: ZARNOVITZ, V. (ed.). **Economic forecasts and expectations: analysis of forecasting behavior and performance**. National Bureau of Economic Research, 1969. p. 3-46.
- MURPHY, K. P. **Machine learning: a probabilistic perspective**. Massachusetts: MIT press, 2012.
- NELSON, B. D. Conditional heteroskedasticity in asset returns: a new approach. **Econometrica**, v. 59, p. 347–370, 1991.
- NEWBOLD, P.; GRANGER, C. W. Experience with forecasting univariate time series and the combination of forecasts. **Journal of the Royal Statistical Society: Series A (General)**, v. 137, n. 2, p. 131-146, 1974.
- NEWBY, W. K.; WEST, K. D. Hypothesis testing with efficient method of moments estimation. **International Economic Review**, p. 777-787, 1987.
- NOWOTARSKI, J.; RAVIV, E.; TRÜCK, S.; WERON, R. An empirical comparison of alternative schemes for combining electricity spot price forecasts. **Energy Economics**, v. 46, p. 395-412, 2014.
- NYSTRUP, P.; MADSEN, H.; LINDSTRÖM, E. Long memory of financial time series and hidden Markov models with time-varying parameters. **Journal of Forecasting**, v. 36, n. 8, p. 989-1002, 2017.
- OJAGVERDIYEVA, S.; PRYSYAZHNYUK, Y. **Testing the Black-Litterman Model Sensitivity of Weight Vector to Variance of Views**. Dissertação (Mestrado em Economia) -, -Lund University, Malmo, 2011.
- PALM, F. C.; ZELLNER, A. To combine or not to combine? Issues of combining forecasts. **Journal of Forecasting**, v. 11, n. 8, p. 687-701, 1992.

- PETTENUZZO, D.; TIMMERMANN, A. **Internet Appendix for “Model Instability and Forecasting Performance”**. 2015. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169207015000418#s0095>. Acesso em: 20 abr. 2022.
- PRIMICERI, G. E. Time varying structural vector autoregressions and monetary policy. **The Review of Economic Studies**, v. 72, n. 3, p. 821-852, 2005.
- PHILLIPS, P. C.; PERRON, P. Testing for a unit root in time series regression. **Biometrika**, v. 75, n. 2, p. 335-346, 1988.
- REDELL, N. **ForecastML**: time Series Forecasting with Machine Learning Methods. R package version 0.8.0. 2020. Disponível em: <https://CRAN.R-project.org/package=forecastML>. Acesso em: 15 fev. 2021.
- SAMUELS, J.; SEKKEL, R. Model confidence sets and forecast combination. **International Journal of Forecasting**, v. 33, p. 48-60, 2017.
- SATCHELL, S. **Forecasting expected returns in the financial markets**. London: Elsevier, 2007.
- SHEPHARD, Neil. **Stochastic Volatility**: selected readings. Oxford: Oxford University Press, 2005.
- STOCK, J. H.; WATSON, M. W. Evidence on structural instability in macroeconomic time series relations. **Journal of Business & Economic Statistics**, v. 14, n. 1, p. 11-30, 1996.
- STOCK, J. H.; WATSON, M. W. Forecasting inflation. **Journal of Monetary Economics**, v. 44, n. 2, p. 293-335, 1999.
- STOCK, J. H.; WATSON, M. W. Forecasting output and inflation: the role of asset prices. **Journal of Economic Literature**, v. 41, n. 3, p. 788-829, 2003.
- STOCK, J. H.; WATSON, M. W. Combination forecasts of output growth in a seven-country data set. **Journal of Forecasting**, v. 23, n. 6, p. 405-430, 2004.
- STOCK, J. H.; WATSON, M. W. Why has US inflation become harder to forecast? **Journal of Money, Credit and banking**, v. 39, p. 3-33, 2007.
- ROSSI, B.; INOUE, A. Out-of-sample forecast tests robust to the choice of window size. **Journal of Business and Economic Statistics**, v.30, n.3, p.432–453, 2012.
- ROSSI, B.; SEKHPOSYAN, T. Forecast rationality tests in the presence of instabilities, with applications to Federal Reserve and survey forecasts. **Journal of Applied Econometrics**, v. 31, n. 3, p. 507–532, 2016.
- ROSSI, B.; SOUPRE, M. Implementing tests for forecast evaluation in the presence of instabilities. **The Stata Journal**, v. 17, n. 4, p. 850-865, 2017.
- RUMLER, F.; VALDERRAMA, M. T. Comparing the New Keynesian Phillips Curve with time series models to forecast inflation. **The North American Journal of Economics and Finance**, v. 21, n. 2, p. 126-144, 2010.
- SATCHELL, S.; SCOWCROFT, A. A demystification of the Black–Litterman model: managing quantitative and traditional construction. **Journal of Asset Management**, v. 1, p. 138–150, 2000.
- TIBSHIRANI, R. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 58, n. 1, p. 267-288, 1996.
- TIMMERMANN, A. Forecast combinations. **Handbook of economic forecasting**, v. 1, p. 135-196, 2006.
- THOMSON, M. E. et al. Combining forecasts: performance and coherence. **International Journal of Forecasting**, v. 35, n. 2, p. 474-484, 2019.

- TRICHES, D.; FEIJÓ, F. T. Uma estimação da Curva de Phillips híbrida para o Brasil no regime de metas de inflação. **Economia Aplicada**, v. 21, n. 1, p. 29-43, 2017.
- WANG, H.; LI, G.; TSAI, C. L. Regression coefficient and autoregressive order shrinkage and selection via the lasso. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 69, n. 1, p. 63-78, 2007.
- WEISS, C. E.; RAVIV, E.; ROETZER, G. Forecast Combinations in R using the ForecastComb Package. **R Journal**, v. 10, n. 2, 2018.
- WEST, K. D.; MCCracken, M. W. Regression-based tests of predictive ability. **International Economic Review**, v. 39, p. 817–840, 1998.
- WRIGHT, J. H. Forecasting US inflation by Bayesian model averaging. **Journal of Forecasting**, v. 28, n. 2, p. 131–144, 2009.
- YOSHIHIRO. **A formação da expectativa inflacionária no regime monetário de metas de inflação e a credibilidade do Banco Central do Brasil**. 2005. Tese (Doutorado em Economia) - Escola de Economia de São Paulo-EESP, São Paulo, 2005.
- ZHANG, B. Real-time inflation forecast combination for time-varying coefficient models. **Journal of Forecasting**, v. 38, n. 3, p. 175-191, 2019.
- ZOU, H.; HASTIE, T.; TIBSHIRANI, R. On the “degrees of freedom” of the lasso. **Annals of Statistics**, v.35, n.5, p. 2173-2192, 2007.

ANEXO - Teste de raiz unitária: Jan/1996 – Dez/2014 (dentro da amostra)

Variável	Nível						Primeira diferença					
	ADF		PP		KPSS		ADF		PP		KPSS	
	Stat.	C.V (5%)	Stat.	C.V (5%)	Stat.	C.V (5%)	Stat.	C.V (5%)	Stat.	C.V (5%)	Stat.	C.V (5%)
IPCA	-7.97	-2.87	-7.95	-2.87	0.16	0.46	-10.90	-2.87	-10.90	-2.87	0.09	0.46
IPCA-15	-7.30	-2.87	-7.35	-2.87	0.20	0.46	-14.01	-2.87	-23.21	-2.87	0.13	0.46
IGP-DI	-6.79	-2.87	-6.86	-2.87	0.32	0.46	-14.15	-2.87	-23.57	-2.87	0.05	0.46
IGP-M	-6.74	-2.87	-6.36	-2.87	0.33	0.46	-13.57	-2.87	-22.81	-2.87	0.07	0.46
IGP-10	-6.87	-2.87	-6.47	-2.87	0.33	0.46	-13.77	-2.87	-20.81	-2.87	0.05	0.46
IPC-Br	2.46	-2.87	4.00	-2.87	1.75	0.46	-5.51	-2.82	-47.23	-2.87	0.17	0.46
IPC-S	-5.42	-2.87	-5.38	-2.87	0.16	0.46	-11.40	-2.87	-21.66	-2.87	0.06	0.46
IPC-Fipe	2.61	-2.85	-1.23	-2.87	1.93	0.46	-10.58	-2.90	-9.92	-2.87	0.10	0.46
INPC	2.24	-2.93	4.00	-2.87	1.97	0.46	-5.35	-2.93	-50.31	-2.87	0.18	0.46
PIB	2.49	-2.87	3.67	-2.87	1.93	0.46	-5.25	-2.87	-51.34	-2.87	0.16	0.46
hiato	2.46	-2.90	3.45	-2.87	2.04	0.46	-5.20	-2.82	-47.74	-2.87	0.15	0.46
IBOV	-1.38	-2.87	-1.29	-2.87	1.79	0.46	-10.90	-2.87	-10.90	-2.87	0.09	0.46
Petr	-1.54	-2.87	-1.62	-2.87	1.82	0.46	-12.99	-2.87	-12.99	-2.87	0.12	0.46
M0	2.15	-2.87	2.77	-2.87	1.94	0.46	-3.67	-2.87	-25.97	-2.87	0.66	0.46
M1	0.56	-2.87	0.91	-2.87	1.98	0.46	-2.97	-2.87	-20.32	-2.87	0.39	0.46
M2	2.45	-2.87	5.27	-2.87	1.82	0.46	-2.80	-2.87	-14.32	-2.87	1.53	0.46
M3	13.64	-2.87	10.86	-2.87	1.83	0.46	-2.08	-2.87	-9.74	-2.87	1.90	0.46
M4	14.11	-2.87	12.55	-2.87	1.85	0.46	-1.92	-2.87	-9.96	-2.87	1.99	0.46
Over Selic	-1.77	-1.94	-1.75	-2.87	1.69	0.46	-17.51	-2.87	-20.62	-2.87	0.13	0.46
Selic.copom	-1.95	-1.94	-1.84	-2.87	1.64	0.46	-16.63	-2.93	-20.62	-2.87	0.13	0.46
Desemp	-0.56	-2.87	-0.36	-2.87	1.39	0.46	-3.58	-2.87	-7.74	-2.87	0.46	0.46
Horas	-2.06	-2.87	-3.57	-2.87	1.09	0.46	-2.89	-2.87	-20.36	-2.87	0.18	0.46
Onibus	-2.19	-2.87	-2.77	-2.87	1.95	0.46	-15.60	-2.87	-15.86	-2.87	0.37	0.46
Energia	-1.92	-2.87	-1.92	-2.87	1.81	0.46	-15.03	-2.87	-15.05	-2.87	0.48	0.46
IMA.B	-0.03	-2.88	0.24	-2.87	1.37	0.46	-5.44	-2.88	-9.39	-2.87	0.11	0.46
IMA.B-5-	2.07	-2.88	2.18	-2.87	1.38	0.46	-8.52	-2.88	-8.63	-2.87	0.51	0.46
IMA.B-5+	0.06	-2.88	-0.21	-2.87	1.35	0.46	-9.15	-2.88	-9.47	-2.87	0.07	0.46
Sup.Prim	1.92	-2.90	1.72	-2.87	1.76	0.46	-8.88	-2.87	-10.48	-2.87	0.30	0.46
DSLPe	1.88	-2.87	1.58	-2.87	1.96	0.46	-9.65	-2.87	-10.18	-2.87	0.34	0.46
DSLPe	-0.38	-2.87	-0.37	-2.87	1.99	0.46	-15.07	-2.87	-15.07	-2.87	0.03	0.46
e-dolar	-1.93	-2.87	-1.95	-2.87	0.40	0.46	-16.43	-2.87	-16.37	-2.87	0.13	0.46
e-turism	-1.93	-2.87	-1.91	-2.87	0.40	0.46	-16.60	-2.87	-16.37	-2.87	0.13	0.46
e-euro	-1.83	-2.87	-1.78	-2.87	0.86	0.46	-16.39	-2.87	-16.37	-2.87	0.09	0.46

Nota: Dickey Fuller Aumentado (ADF), Phillips Peron (PP) e Kwiatkowski – Phillips – Schmidt – Shin (KPSS)