

Descobrimos Modelos de Previsão para a Inflação Brasileira: Uma Análise a partir de uma Gama Ampla de Indicadores ♦

Anderson Moriya Silva¹

Emerson Fernandes Marçal²

Resumo

Este trabalho visa avaliar o poder preditivo que séries macroeconômicas têm sobre o índice de preços ao consumidor amplo brasileiro (IPCA) utilizando técnicas de séries de tempo. As previsões são realizadas para um horizonte de até 12 períodos à frente e comparadas com um processo autoregressivo como referência. O período vai de janeiro de 2000 até agosto de 2015. Utilizou-se um conjunto amplo de informação de 1 170 séries. Para cada momento e horizonte de tempo selecionou-se um novo modelo utilizando o algoritmo Autometrics desenvolvido por Hendry e Doornik (2014). O desempenho preditivo dos modelos foi comparado utilizando o Model Confidence Set, desenvolvido por Hansen, Lunde and Nason (2010). Os resultados sugerem que há ganhos expressivos de previsão principalmente para os horizontes mais longos.

Palavras-Chave

Inflação. Previsão. Seleção de modelos. Autometrics. Model Confidence Set.

Abstract

This work evaluates the prediction capabilities of econometric time series models based on macroeconomics indicators for Brazilian consumer price index (IPCA). We run a pseudo real time prediction exercise with twelve steps ahead horizon. Predictions are compared with univariate models such as the first order autoregressive model among others. The sample period goes from January 2000 to August 2015. We evaluated over 1 170 different economic variable for each forecast period, searching for the best predictor set in each point in time using Autometrics algorithm as model selector. Models' performance is compared using Model Confidence Set, developed by Hansen, Lunde and Nason (2010). Our results suggest possible gains in predictions that use a high number of indicators particularly at longer horizons.

♦ Os autores agradecem a Marcelo Kfoury e Pedro Luiz Valls Pereira pelos comentários e sugestões feitas a uma versão preliminar deste trabalho. Os erros remanescentes são de nossa responsabilidade.

¹ Cientista da Computação – Microsoft

Endereço: Av das Nações Unidas, 12901 – 31º andar – Brooklin Novo – São Paulo/SP

CEP: 04578-910 – E-mail: anderson.moriya@microsoft.com – <https://orcid.org/0000-0002-5497-0841>

² Professor – Escola de Economia de São Paulo (EESP/FGV)

Endereço: Rua Itapeva, 286 - 10º andar – Bela Vista – São Paulo/SP

CEP: 01332-000 – E-mail: emerson.marcal@fgv.br – <https://orcid.org/0000-0002-0841-5644>

Recebido: 23/01/2017. Aceite: 21/08/2017.



Esta obra está licenciada com uma Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Keywords

Inflation. Forecasting. Model Selection. Autometrics. Model Confidence Set.

JEL Classification

C53. C55. E37.

1. Introdução e Motivação

O Brasil, após décadas de inflação alta, conseguiu estabilizar sua economia e passou a apresentar médias de inflação mais próximas da média mundial. Dentro deste processo o país adotou o regime de metas de inflação a partir de março de 1999. O Banco Central é o responsável, embora sem independência, pelo controle de preços na economia, comprometendo-se a tomar as medidas necessárias para atingir uma meta preestabelecida. Como a meta é conhecida e amplamente divulgada e o Banco Central deve desviar da mesma apenas de forma esporádica, modelos bem simples de previsão tal como um modelo autoregressivo devem fornecer previsões acuradas, sendo difícil batê-los sistematicamente. Isto já acontece em países com tradição inflacionária mais baixa e de boa gestão macroeconômica.

A proposta do presente trabalho é explorar um grande número de variáveis econômicas tendo em vista descobrir modelos de previsão baseados em indicadores para a inflação brasileira. A inovação deste trabalho, em relação a outros previamente desenvolvidos para previsão da inflação brasileira em um ambiente com grande quantidade de dados, é que ao invés de utilizar modelos baseados em fatores como proposto por Bernanke, Boivin e Eliaz (2005), será aplicado o algoritmo seleção automática chamado *Autometrics*, que segue a abordagem geral para o específico. O algoritmo foi desenvolvido por Hendry e Doornik (Doornik (2007, 2009a), Hendry e Doornik (2013a)). Dado uma determinada amostra, através de testes de especificação e significância estatística, o algoritmo elimina as variáveis não relevantes procurando aproximar o processo gerador de dados local (PGDL). O objetivo desse tipo de abordagem é descobrir o modelo de forma empírica que seja congruente e abrangente.

Para ranquear os modelos em termos de poder preditivo iremos comparar os resultados das previsões com os dos modelos autoregressivos vistos como *benchmark*. A metodologia de comparação a ser adotada é o *Model*

Confidence Set (MCS), introduzido por Hansen, Lunde e Nason (2010), que compara os modelos de previsão bem como estabelece uma probabilidade de o modelo fazer parte do conjunto de melhores modelos para previsão.

Coleta-se evidência de que modelos baseados em uma gama ampla de indicadores acoplados a algoritmos de seleção automática apresentam ganhos de precisão em previsões para inflação brasileira em diversos horizontes. Os ganhos são maiores para horizontes mais longos próximos a um ano.

O trabalho está organizado em sete seções. A primeira seção é dada por esta introdução. A segunda consiste na revisão de literatura na qual os algoritmos *Autometrics* e o *Model Confidence Set* são descritos. Na terceira seção é apresentada a base de dados utilizada. Na quarta seção, apresentam-se os modelos utilizados. Na quinta seção os resultados são discutidos. Na sexta seção, as limitações e extensões do trabalho são brevemente discutidas. Por fim as conclusões são apresentadas.

2. Revisão Teórica

Existem diversas abordagens para a previsão do nível de preços na economia. Entre elas podemos citar as mais clássicas como modelos baseados na curva de Philips e os Vetores Autoregressivos (VAR). A curva de Philips define a relação de curto prazo entre a taxa de inflação e o desemprego. Quando se tem uma taxa de desemprego abaixo da taxa de equilíbrio, espera-se uma elevação na inflação sendo que o equilíbrio entre o desemprego e a taxa de inflação é chamada de NAIRU.³ Segundo Atekeson e Ohanian (2001), modelos baseados na curva de Philips têm sido amplamente utilizados para prever a inflação na literatura acadêmica. Apesar de ser bastante difundido, não existem conclusões definitivas com relação a sua efetividade como pode ser visto nos estudos de Cecchetti, Chu and Steindel (2000), Kozicki (2001) e Stock e Watson (2002a).

Além do desemprego existem outros componentes que também afetam o nível de preços agregado como, por exemplo, poupança, decisão de investimento das famílias e firmas, gastos do governo, taxa de câmbio, entre

³ Non Accelerating Inflation Rate of Unemployment.

outros. Nessa linha, os modelos VAR, conforme descreve Sims (1980), têm se tornado uma alternativa interessante e com grande apelo para previsão de inflação. Um dos fatores que levou a sua popularidade é que todas as variáveis são consideradas endógenas e o modelo é estimado em sua forma reduzida. O VAR se tornou bastante utilizado nos bancos centrais em estudos sobre a transmissão da política monetária. Um dos problemas dessa abordagem conforme Bernanke *et al.* (2005) é que o número de variáveis utilizadas nesses modelos é pequeno se comparado com a quantidade de informações disponíveis, o que pode gerar problemas de modelos com variáveis omitidas.

Diversas metodologias têm sido desenvolvidas com o propósito de obter previsões mais precisas em um ambiente em que o número de dados e variáveis disponíveis cresce exponencialmente. Stock e Watson (2006) descrevem algumas abordagens de previsão utilizando um grande conjunto de variáveis e informações. Segundo Castle, Clements e Hendry (2011) existem três principais tradições em previsão econômica, sendo que duas delas utilizam um grande conjunto de informações baseado em indicadores ou fatores. Como membros do primeiro grupo podemos citar os modelos de equilíbrio geral estocástico e dinâmicos (*dynamic stochastic general equilibrium - DSGE*) Smets e Wouter (2003) e modelos Vetoriais e autorregressivos globais (Global VAR – GVAR - Dees, di Mauro, Pesaran and Smith (2007), Pesaran, Schuermann e Smith (2009) e Ericsson (2010). A segunda abordagem inclui modelos baseados em índices de difusão e modelos fatoriais que utilizam um grande número de variáveis econômicas (Stock e Watson, 1989, 1999 e 2009). Já a terceira refere-se a modelos tradicionais de econometria de séries de tempo como Box and Jenkins (1970) e frequentemente aparecem como dominante em competições de previsão conforme Makridakis, Andersen, Carbone, Fildes *et al.* (1982) e Makridakis e Hibon (2000).

2.1. Modelagem Empírica em Economia

A abordagem de seleção automática de modelos proposta por Hendry e Doornik (2013a) parte da noção de que o processo gerador dos dados (PGD) deve ser descoberto e aproximado por um determinado conjunto de variáveis. Como o processo gerador dos dados é extremamente complexo, o econometrista em geral faz uma aproximação local, condicionada a um con-

junto de informação mais restrito. Isto é denominado de processo gerador dos dados local (PGDL). Esta abordagem é um refinamento importante da tradição de seleção de modelos baseadas na abordagem geral para o específico (GETS).

Em princípio, pode existir um grande número de variáveis candidatas a explicar determinado fenômeno assim como o próprio passado da série. Após a definição do conjunto de possíveis candidatas a explicar o fenômeno em questão, o objetivo passa a ser identificar, dentro do universo de múltiplos regressores, aqueles que são realmente relevantes. Esse procedimento gera o subconjunto x_t^* . O propósito é procurar o subconjunto de variáveis que seja uma boa aproximação do processo local gerador de dados.

Através da abordagem adotada em Doornik e Hendry (2014), existem alguns aspectos principais para o processo de descoberta na modelagem empírica em Economia. O primeiro passo é especificar y_t , a variável alvo e o conjunto N de variáveis candidatas a explicá-la. Nesse momento, a utilização da teoria econômica previamente estabelecida, à luz de evidência passada, tem um papel bastante importante. Como a intenção é de descoberta empírica, também não se deve restringir o número de variáveis a um universo muito reduzido, uma vez que dessa forma pouca nova informação poderá ser extraída. O segundo passo envolve definir a variável x de forma mais genérica, considerando suas relações dinâmicas e formas funcionais em um modelo mais geral com $x_1 \dots x_{t-s}$ e os termos não lineares $g(x_t) \dots g(x_{t-s})$, na qual a maior defasagem é $s \geq 0$ períodos e $g(\cdot)$ denota as funções apropriadas do subconjunto de variáveis x , com um conjunto de T indicadores para quebras e *outliers* ilustrado pela variável indicadora $1_{\{i=T\}}$, ou por mudanças de interceptos, denotados por $1_{\{i<T\}}$, gerando o Modelo Irrestrito Geral (MIG):

$$y_t = \sum_{j=0}^s \beta'_j x_{t-j} + \sum_{j=0}^s k'_j g(x_{t-j}) + \sum_{j=0}^s \lambda_j y_{t-j} + \sum_{i=1}^T \delta_i 1_{\{i=t\}} + \sum_{i=1}^{T-1} \varphi_i 1_{\{t \leq i\}} + \varepsilon_t \quad (1)$$

O objetivo é descobrir as variáveis associadas e as formas funcionais que forem relevantes representadas por uma função $g(x_{t-i}^*)$ que inclui as defasagens e indicadores. Esta função pode ser aproximada por termos polinomiais de várias ordens. Ver Hendry e Doornik (2015) para maiores detalhes na formulação do MIG.

Considerando o vetor d que contém um subconjunto das *dummies* indicadores ($1_{\{i=t\}}$) e das *dummies* de intervenção ($1_{\{t \leq i\}}$) não excluídas durante a seleção e as variáveis com asteriscos como um subconjunto de variáveis relevantes descobertas a partir do conjunto geral sem asterisco, a representação do modelo final congruente e parcimoniosamente abrangente é:

$$y_t = \sum_{j=0}^{s^*} \lambda_j^* y_{t-j} + \sum_{i=1}^{s^*} (\beta_i^*)' x_{t-i}^* + \sum_{i=1}^{s^*} (k_i^*)' g(x_{t-i}^*) + \gamma' d_t + v_t \quad (2)$$

Na qual v_t é o componente inexplicado ou ruído branco. Como pode-se observar no MIG é possível atribuir um *outlier* para cada observação o que naturalmente leva a uma situação em que o número de regressores é maior que observações ($N > T$).

O procedimento automático de seleção pode ser considerado bem-sucedido quando consegue localizar uma representação congruente e parcimoniosa do modelo a ser reduzido. Além de ser automático, o procedimento deve ter uma baixa probabilidade de reter variáveis irrelevantes bem como probabilidade controlada de manter as variáveis relevantes como se o PGDL estivesse sido especificado diretamente. Nesse contexto, existem evidências que o *Autometrics* possui essas propriedades dentro de certas condições.

Este trabalho não tem por objetivo fazer uma descrição exaustiva do algoritmo *Autometrics* e de suas propriedades estatísticas. Uma descrição detalhada do mesmo pode ser encontrada em Doornik & Hendry (2014 e 2015)

2.2. Teoria da Redução

De acordo com Hendry (1987, 1995a, 2009), o processo gerador dos dados consiste na distribuição conjunta de todas as variáveis na economia sendo formado por inúmeros agentes econômicos. Criar teorias ou modelos sobre o PGD é inviável, uma vez que este possui inúmeras dimensões, é extremamente heterogêneo e não estacionário; desta forma, o problema a ser analisado deve ser reduzido a um tamanho no qual seja possível sua manipulação. Isso é alcançado reduzindo o problema ao espaço das variáveis $\{x_t\}$, o que é assim chamado de processo gerador de dados local (PGDL).

O PGDL é a distribuição conjunta de probabilidade $D_X(x_1 \dots x_T | \Theta, X_0)$, na qual X_0 é o conjunto inicial de informação e o parâmetro $\theta \in \Theta$ pode variar no tempo. Considerando $D_X(\cdot)$ como sendo o processo gerador de dados para $\{x_t\}$, uma vez conhecido o PGDL, é possível obter o comportamento provável dos dados para $\{x_t\}$ que deve desviar dos dados observados apenas por um ruído aleatório. Portanto, nada melhor pode ser feito além de conhecer o $D_X(\cdot)$, tornando a busca pelo PGDL o principal objetivo a ser alcançado pela seleção de modelos.

Como o objetivo da modelagem empírica é descobrir as propriedades de $D_X(\cdot)$ através de todas as informações disponíveis, conhecimento teórico, relevância histórica entre outros, criar um modelo inicial que liste todas essas informações e que principalmente contenha a informação sobre o PGDL é crucial para o sucesso da seleção de modelos. Nesse contexto, o Modelo Irrestrito Geral é o ponto de partida para esse processo. As propriedades de consistência dos modelos de seleção que partem do geral para o específico, supõem que o MIG contenha o PGDL como caso particular conforme Campos *et al.* (2003). Como isso é improvável de ser alcançado em trabalhos empíricos, espera-se que ele seja ao menos uma boa aproximação do PGDL. White (1990) mostra que é possível selecionar corretamente os modelos utilizando a abordagem geral para o específico.

As variáveis não significativas são eliminadas do MIG com o objetivo de reduzir sua complexidade. A cada redução, testes de diagnóstico são executados a fim de verificar se a transformação continua garantindo também a congruência do modelo final. Sucessivos modelos congruentes devem abranger os anteriores de forma parcimoniosa e progressiva.

2.3. Seleção de Modelos com Base na Regra “De um Corte Só”

Um dos problemas encontrados na seleção de modelos baseado em dados é a quantidade de modelos possíveis, 2^N , a serem avaliados. Utilizar métodos tradicionais como critério de informação Akaike (1973) ou similares nessa quantidade de possibilidades se torna inviável, posto que o número de combinações cresce exponencialmente com o número de variáveis.

Considerando um caso particular na qual todos os regressores são mutuamente ortogonais, independente dos erros e os parâmetros do PGDL são

constantes e $N \gg T$, é possível demonstrar que não é necessário analisar ou avaliar 2^N possíveis modelos para se obter uma boa aproximação do PGDL. Apenas uma única decisão é necessária, o que é conhecido como “*decisão de um corte só*”. As seleções de modelos baseados na abordagem geral para o específico (GETS) podem ser vistos como uma generalização na qual os regressores não são perfeitamente ortogonais e o número de variáveis é grande, conforme Hendry, Castle e Doornik (2011).

Partindo modelo abaixo, a regra de “um corte só” pode ser implementada da seguinte forma:

$$y_t = \sum_{i=1}^N \beta_i z_{i,t} + \varepsilon_t \text{ sendo } \varepsilon_t \sim \text{IN}[0, \sigma_\varepsilon^2] \quad (3)$$

Na qual $T^{-1} \sum_{t=1}^T z_{i,t} z_{j,t} = \lambda_{\lambda i, i}$ para $i = j$ e $0 \forall i \neq j$. Após a estimação, ordenar as estatísticas t^2 , testando a hipótese nula que $H_0: \beta_k = 0$ conforme:

$$t_{(1)}^2 \geq t_{(2)}^2 \geq \dots \geq t_{(N)}^2 \quad (4)$$

A linha de corte m entre as variáveis excluídas ou não é dada por:

$$t_{(m)}^2 \geq c_{(\alpha)}^2 \geq t_{(m-1)}^2 \quad (5)$$

Portanto as variáveis com maior valor t^2 são retidas e as outras excluídas. Esse processo se dá com apenas uma única decisão. A Tabela 1 lista as probabilidades de rejeição da nula quando efetuando testes t para N regressores irrelevantes ao nível de significância α .

Tabela 1 – regra de “um corte só”. Probabilidade de rejeição da nula

Evento	Probabilidade	No. Rejeitado
$P(t_i < c_\alpha, \forall i = 1, \dots, N)$	$(1 - \alpha)^N$	0
$P(t_i \geq c_\alpha \mid t_j < c_\alpha, \forall i \neq j)$	$N\alpha(1 - \alpha)^{N-1}$	1
$P(t_i \geq c_\alpha, t_k \geq c_\alpha \mid t_j < c_\alpha, \forall j \neq i, k)$	$\frac{1}{2}N(N-1)\alpha^2(1 - \alpha)^{N-2}$	2
⋮	⋮	⋮
$P(t_i < c_\alpha \mid t_j \geq c_\alpha, \forall i \neq j)$	$N\alpha^{N-1}(1 - \alpha)$	N-1
$P(t_i \geq c_\alpha, \forall i = 1, \dots, N)$	α^N	N

A Tabela 1 corresponde a expansão de $(1 - \alpha)^N$, sendo α^N a probabilidade de que todos N coeficientes sejam significativos. A média de retenção é dada pela Equação 6.

$$k = \sum_{i=0}^N i \frac{N!}{i!(N-i)!} \alpha^i (1 - \alpha)^{N-1} = N\alpha \quad (6)$$

Desta forma, a seleção de 1000 variáveis com um $\alpha = 0,001$ gera um k de 1, ou seja, existe a probabilidade de que nessa seleção ao menos 1 regressor seja mantido mesmo sendo todos irrelevantes. Esse conceito é denominado como *gauge* e denota a frequência de retenção de variáveis irrelevantes durante o procedimento de seleção. A contrapartida de se dispor de valores de α extremamente pequenos, afim de evitar qualquer possibilidade de retenção de regressores irrelevantes é a perda de poder de detecção de regressores relevantes. Essa característica é denominada de *potency* e é a taxa de retenção sobre as variáveis realmente relevantes.

2.4. Autometrics

O *Autometrics* teve sua origem na chamada abordagem da *London School of Economics* para a Econometria. Baseia-se na modelagem geral para específico que consiste em partir de um modelo irrestrito geral e utilizar procedimentos estatísticos para avaliar a relevância das variáveis com objetivo de eliminar as não pertinentes. Diferentemente da abordagem da seleção de modelos com uma decisão, a cada passo de redução testes de erro de especificação são efetuados a fim verificar a congruência dos modelos após as eliminações. O objetivo desse procedimento é localizar o modelo abrangente e parcimonioso que seja uma boa representação do processo gerador de dados local. De acordo com Hendry e Nielsen (2007), um modelo é considerado congruente com a informação disponível quando satisfaz os testes de erro de especificação para (i) heterocedasticidade, autocorrelação e não normalidade, (ii) falha na hipótese de exogeneidade fraca, (iii) constância nos parâmetros.

A abordagem geral para específico tem seu desenvolvimento situado no desenvolvimento da teoria de redução. Campos, Ericson e Hendry (2005b) discutem esta evolução detalhadamente. Os avanços computacionais das

últimas décadas abriram espaço para o desenvolvimento de algoritmos automáticos de seleção. Os trabalhos de Hoover e Perez (1999) foram pioneiros nesta área. Analisando múltiplos caminhos de redução, os autores conseguiram reduzir a taxa de retenção de variáveis irrelevantes e aumentar a de variáveis relevantes utilizando o algoritmo GETS proposto por eles.

Hendry e Krolzig (1999) propuseram melhorias ao algoritmo GETs de Hoover e Perez através do *PCGets* que fora implementado no software econométrico *Oxmetrics*. Através de simulações de Monte Carlo, eles analisaram a probabilidade do *PcGets* obter o processo gerados dos dados e conseguiram bons resultados com o *gauge* perto do nominal sem grandes perdas de *potency*, relativo ao teórico. Apesar do *Autometrics* ter como inspiração o *PcGets*, ele possui diversas melhorias como a busca em árvores, não supõe ortogonalidade das variáveis, permite que N seja bem maior que T , lida com perfeita colinearidade gerada pelas *dummies* de saturação. Ericsson e Kamin (2009) demonstraram a superioridade do *Autometrics* quanto a capacidade de obter modelos mais parcimoniosos e abrangentes do que o *PCGets*.

Para evitar problemas referentes a *outliers* e quebras estruturais, o *Autometrics* utiliza o *impulse-indicator saturation (IIS)*, conforme Hendry, Johansen e Santos (2008) e *step-indicator saturation (SIS)*, Castle, Doornik, Hendry e Pretis (2013b). A estratégia do IIS consiste em adicionar *dummies* a cada observação e no espírito do *Autometrics* verificar quais que são mantidas nos modelos terminais finais. No caso do SIS, são adicionadas as *dummies* do início da amostra até a n -ésima observação. O objetivo é capturar os períodos em que houve quebras estruturais permanentes. Para tratar a situação de $N > T$, o algoritmo quebra o conjunto de variáveis em sub-blocos que contém um número de regressores menor que o total das observações.

Existem diversos estudos que comprovam o desempenho do *Autometrics* com outras abordagens de seleção automática de modelos como Hendry, Castle e Doornik (2011) que demonstram a capacidade do algoritmo de trabalhar com variáveis não ortogonais e correlação alta e com $N > T$.

2.5. Model Confidence Set

O *Model Confidence Set* (MCS) tem como objetivo efetuar a comparação de modelos de previsão utilizando a informação disponível para determinar quais são os melhores dentro de um determinado nível de significância. Introduzido por Hansen *et al.* (2010), o *Model Confidence Set* permite determinar um conjunto de modelos M^* , que contém os melhores modelos de um conjunto M^0 modelos, sendo que os melhores são definidos através de um intervalo de confiança estipulado pelo pesquisador. O MCS baseia-se em uma regra de eliminação e_M , e em um teste de equivalência δ_M . Estes são aplicados ao conjunto $M = M^0$. Se houver rejeição da hipótese de equivalência, a regra de eliminação será utilizada para remover um modelo M com desempenho inferior. Esse procedimento é repetido até que a hipótese de equivalência não seja rejeitada a um determinado nível de significância. É neste ponto que são definidos o conjunto de melhores modelos.

Uma das vantagens dessa abordagem consiste em não escolher um *benchmark*. No contexto de comparação de previsões o problema fica evidente quando efetuamos testes de capacidade de previsão superior (*superior predictive ability* - SPA White (2000b), Hansen (2005) e Romano e Wolf (2005)). Diferentemente do teste de realidade (*reality check for data snooping*) de White (2000b) e o teste de SPA de Hansen (2005) que precisam de um *benchmark* para comparação, o MCS compara os modelos entre si o que é particularmente interessante quando a escolha do *benchmark* não é clara.

Os modelos são ranqueados através dos seus p-valores. Utilizando o nível de significância, α , é possível garantir que $\lim_{n \rightarrow \infty} P(M^* \subset \widehat{M}_{1-\alpha}^*) \geq 1 - \alpha$, e se houver algum objeto com resultado mais expressivo, temos $\lim_{n \rightarrow \infty} P(M^* = \widehat{M}_{1-\alpha}^*) = 1$. Qualquer objeto que tiver um p-valor menor que α não está no subconjunto dos melhores modelos M^* , ou seja, para cada objeto $i \in M^0$, o p-valor do MCS, \hat{p}_i , é o limite em que $i \in \widehat{M}_{1-\alpha}^*$ se e somente se $\hat{p}_i \geq \alpha$.

É importante notar que embora o MCS mantenha em M^* os melhores modelos assintoticamente, em amostra finita é possível que sejam mantidos modelos não tão bons dado o baixo poder para eliminá-los. Como o MCS tem o mesmo princípio de inferência equivalente a testes de significância, um modelo só é descartado quando este é estatisticamente inferior a

outro, permanecendo no conjunto M^* até então. Desta forma não são todos os modelos que permanecem no MCS que podem ser considerados bons.

Outro ponto interessante do MCS é o reconhecimento da limitação dos dados, ou seja, com dados mais informativos é possível classificar os modelos entre os melhores e piores. Dados menos informativos tornam o processo de distinção entre os modelos mais difícil e o MCS pode retornar todos os modelos como sendo equivalentes. Essa característica não é levada em consideração pela maioria das funções de perda comumente usadas para comparações de previsões de modelos.

Em Hansen, Lunde e Nason (2010), os autores reavaliam as projeções de inflação de Stock e Watson (1999) utilizando o MCS. Stock e Watson (1999) concluem que modelos baseados na curva de Philips, se bem especificados, são as melhores ferramentas para previsão de inflação. Já Atkeson e Ohanian (2001) concluem que tais modelos dificilmente são melhores que um simples (*no-change forecast*). Hansen, Lunde e Nason (2010), verificam que em períodos de maior volatilidade como na sub-amostra de 1970:M1 – 1983:M12, fica claro quais modelos são melhores, já em períodos menos voláteis como 1984:M1 – 1996:M9, a série não é informativa o suficiente, e o MCS considerou equivalentes quase todos os modelos avaliados por Stock e Watson.

Segundo Hansen, Lunde e Nason (2010), a pergunta sobre qual é o melhor modelo de previsão de inflação para os Estados Unidos nos últimos 35 anos continua uma questão não respondida, uma vez que dada a insuficiente informação contida nos dados não é possível distinguir entre os melhores e piores modelos.

3. Dados e Período Analisado

Os dados utilizados neste trabalho foram obtidos através do Módulo Público do Sistema Gerenciador de Séries Temporais (SGS) disponibilizado pelo Banco Central do Brasil. A variável a ser analisada, endógena no nosso modelo de previsão, é a variação mensal da inflação agregada brasileira registrada pelo IBGE sob o código de IPCA - Índice nacional de preços ao consumidor-amplo, número 433 no sistema de série temporal do

Banco Central do Brasil. As variáveis explicativas também têm periodicidade mensal e foram obtidas através do *WebService* disponibilizado pelo Banco Central. Ao todo foram obtidas 1170 variáveis explicativas candidatas para o período de janeiro de 2000 a setembro de 2015.

Aplicou-se o logaritmo neperiano nas séries que não possuíam números negativos e a transformação $\ln(1 + \frac{x}{100})$ para as séries que são disponibilizadas em variações ou taxas seguindo procedimento adotado por Figueiredo (2010). Também para evitar problemas associados a não estacionaridade das séries adotamos o procedimento sugerido por Enders (2010) aplicando teste de Dickey-Fuller Aumentado (ADF) (1979) a fim de verificar a existência de raiz unitária. Caso identificado, foi tirado a primeira ou segunda diferença para transformá-las em estacionárias.

A amostra utilizada para criar os modelos de previsão através do *Autometrics* iniciou em 2000 com um total de 132 observações em janela móvel, isto é, a cada passo de estimação e previsão a data de início e fim do modelo é deslocada conforme Figura 1. A data inicial de previsão é de novembro de 2011 até setembro de 2015 totalizando 36 previsões de um a doze passos à frente. O tamanho da amostra e quantidade de previsões a serem analisadas foram baseados em Castle, Clements e Hendry (2011).

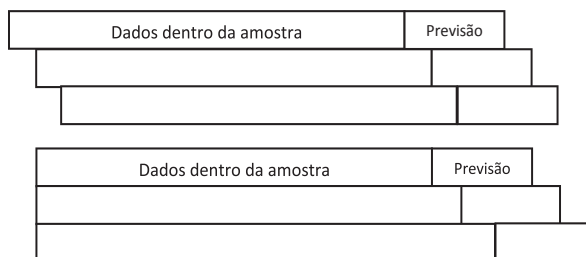


Figura 1 - Janela móvel ou *rolling forward*

Também foi adotado em um dos modelos com múltiplos regressores e em todos os modelos univariados a proposta de janela crescente conforme ilustrado na Figura 2. Uma vez configurada a data inicial dos dados dentro da amostra, esta vai incorporando nova informação assim que as previsões vão avançando no tempo.

O período adotado também levou em consideração o número de séries de tempo novas disponibilizadas pelo Banco Central com o passar do tempo conforme ilustrado na Figura 2. Uma janela de tempo inicial e final de previsão de inflação similar também foi adotada por Figueiredo e Guillén (2013). Nesse trabalho o autor efetuou a previsão de inflação brasileira utilizando FAVAR (*Factor Augmented VAR*) e amostra utilizada foi de junho de 2000 a dezembro de 2012, sendo a avaliação de previsão fora da amostra de janeiro de 2008 a dezembro de 2012 até 24 passos à frente.

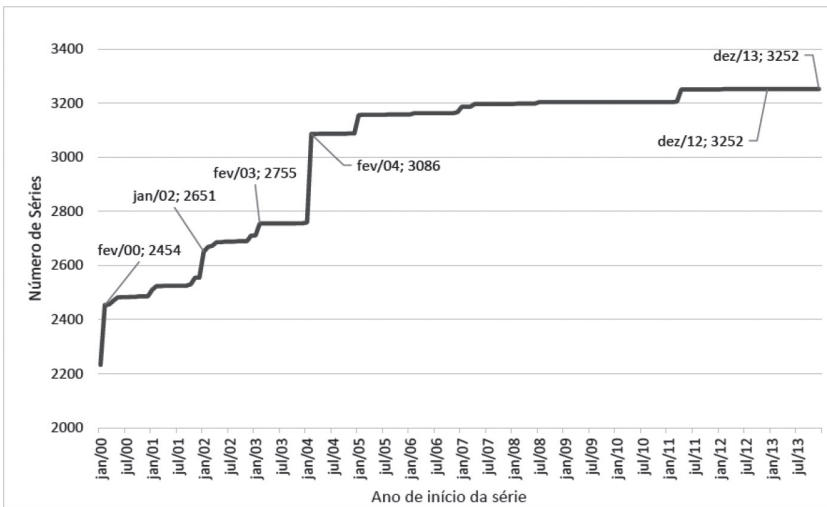


Figura 2 - Números de séries temporais segundo data de início

Fonte: Bader, Koyama e Tsuchida (2014).

4. Modelos de Previsão

4.1. Modelos Multivariados

Os modelos de previsão de um a doze passos à frente foram obtidos através da seleção automática do seguinte MIG:

$$y_t = \sum_{j=J_a+k}^{J_b+k} \lambda_j y_{t-j} + \sum_{i=1}^n \sum_{j=J_a+k}^{J_b+k} \beta_{i,j} z_{i,t-j} + \sum_{l=1}^T \delta_l 1_{\{i=t\}} + \sum_{i=1}^{T-1} \varphi_i 1_{\{t \leq i\}} + \varepsilon_t \quad (7)$$

Na qual y_t a taxa de variação mensal da inflação brasileira após a transformação $\ln(1 + \frac{x}{100})$ e as variáveis candidatas. Como o modelo irá gerar previsões de um a doze passos à frente, k irá variar de 1 a 12 passos. Os valores de J_a e J_b são 0 e 5 respectivamente, ou seja, quando for estimado um passo a frente k será igual a $J_a + k$ e $J_b + k$ 1 e 6 respectivamente. Isso permitirá que o modelo contenha defasagens de um a seis da variável dependente y_t e exógenas. Na iteração seguinte, quando for estimado dois passos à frente, os valores de $J_a + k$ e $J_b + k$ serão dois e sete permitindo a mesma defasagem das variáveis endógenas e exógenas. Esse tipo de formulação possibilita que sejam usadas todas as informações disponíveis até o momento referentes a cada período a ser projetado, um a doze. Também foi estimado um modelo com duas defasagens para cada período a ser projetado, conforme a formulação acima. Configuraram-se os valores de J_a e J_b como 1 e 6 permitindo seis defasagens como no modelo anterior, porém para o período de previsão de um passo à frente, o MIG foi estimado com as defasagens de 2 a 7 respectivamente. O propósito desse modelo é evitar problemas que podem ocorrer por utilizar variáveis que não estavam disponíveis em tempo real.

Foram utilizadas duas estratégias para escolha do tamanho da amostra para os modelos multivariados. A primeira delas foi a janela móvel conhecida como *rolling forward* com tamanho de amostra fixa de $T = 132$, nesse caso quando o período de previsão é deslocado, a data de início da amostra também é deslocada. Uma segunda estratégia adotada foi a de fixar a data de início da amostra e ao deslocar as previsões, também é deslocado o tamanho da amostra. Nesse caso a previsão com maior quantidade de dados teve $T=121$ e menor quantidade $T=90$.

Com relação às variáveis utilizadas, foram feitos dois ensaios. No primeiro deles foram utilizadas todas as variáveis disponíveis (1885). Com essa quantidade de variáveis, utilizando seis defasagens, somando a *dummies* para o procedimento com IIS e SIS, conforme modelagem dada por (7), geram um MIG com 7278 regressores para $T=133$. Para controlar a probabilidade do *Autometrics* reter variáveis irrelevantes no modelo final, configuramos o valor de $\alpha = 0,0001$. Isso gera a probabilidade de que aproximadamente 0,72 regressores não relevantes em média estejam no modelo final. Um segundo ensaio foi feito removendo algumas variáveis que, de acordo com o conhecimento prévio, poderiam ter pouco impacto sobre ele. Foram removidas 377 variáveis referentes a dados desagregados de exportação como por exemplo, Exportações Eslovênia, Exportações Mauricio,

Exportações Guiné, entre outros. Após esse tratamento nos dados foram obtidas 717 variáveis que somando as defasagens, IIS e SIS, geraram um total de 5022 regressores. Para essa simulação também foi utilizado o valor de $\alpha = 0,0001$, gerando a probabilidade de que aproximadamente 0,50 regressores foram retidos sobre a hipótese nula que nenhum regressor é relevante.

4.2. Modelos Univariados

Para comparação com os modelos com múltiplos regressores tendo a definição congruente, parcimoniosa e abrangente dos MIGs acima discriminados, iremos utilizar alguns modelos univariados considerados como *benchmarks*. São eles AR(1), AR(12):

$$\hat{y}_{T+k+h}^{AR(D)} = \hat{\beta}_0 + \hat{\beta}_1 y_{T+k} \quad (8)$$

$$\hat{y}_{T+k+h}^{AR(1)} = \sum_{i=1}^{h-1} \hat{\gamma}_0 \hat{\gamma}_1^i + \hat{\gamma}_1^h y_{T+k} \quad (9)$$

sendo $k = 0, \dots, H - h, h = 1 \dots 12$.

Além desses modelos ingênuos também iremos adicionar um modelo que consiste em um autoregressivo de ordem definida pelo *Autometrics* a cada amostra dada pela janela de projeção, com e sem IIS e SIS. O MIG desses modelos é dado por:

$$y_t = \sum_{j=J_a}^{J_b+k} \lambda_j y_{t-j} + \sum_{l=1}^T \delta_l 1_{\{i=t\}} + \varepsilon_t \quad (10)$$

Sendo $\delta_i = 0$ sem IIS e $\delta_i = 1$ com IIS. Para esses modelos foram utilizados $\alpha = 0,05$.

Tabela 2 – Modelos Estimados e suas configurações

Sigla	Descrição	Tipo	No. de Regressores MIG ¹	α
Auto_1170_GW	Autometrics, 1170 variáveis, IIS, SIS e janela crescente. Fig 3.2	Multivariado	7278	0,0001
Auto_793_GW	Autometrics, 793 variáveis, IIS, SIS e janela crescente. Fig 3.2	Multivariado	5022	0,0001
Auto_793_FW	Autometrics com 1170 variáveis, IIS, SIS e janela móvel Fig 3.1	Multivariado	5020	0,0001
Auto_793_FW_D	Autometrics com 793 variáveis, IIS, SIS e janela móvel Fig 3.1. A defasagem é duas a mais do passo a ser estimado.	Multivariado	5020	0,0001
Auto_793_FW_D_NDS	Autometrics com 793 variáveis, sem utilizar IIS e SIS, janela móvel Fig 3.1. A defasagem é duas a mais do passo a ser estimado	Multivariado	5020	0,0001
AR1 Sazonal	Modelo ingênuo com dummies sazonais, janela crescente	Univariado	-	-
AR12	Modelo ingênuo sem sazonalidade janela crescente	Univariado	-	-
AR1_Raw	Modelo ingênuo, janela crescente	Univariado	-	-
AR12_Auto	Autometrics definindo a defasagem da endógena e janela crescente Fig 3.1	Univariado	24	0,05
AR12_Auto_IIS_SIS	Autometrics definindo a defasagem da endógena, IIS, SIS e janela crescente Fig 3.1	Univariado	198	0,05

Fonte: Elaborado pelos autores.

5. Resultados

A análise de desempenho dos modelos de previsão deu-se através dos resultados acumulados fora da amostra de um a doze passos à frente¹ do período de novembro de 2011 a setembro de 2015. Nos modelos com múltiplo regressores ou nos univariados na qual ocorreu a seleção automática, a cada deslocamento da janela de previsão novas informações são incorporadas e um novo MIG é apresentado ao *Autometrics*, afim buscar

o melhor conjunto de variáveis com toda informação disponível até aquele momento. Para os modelos *ingênuo* como AR(1), AR(1) com sazonalidade, AR(12) não houve seleção automática de variáveis e a cada deslocamento de janela seus parâmetros foram reestimados.

Adotou-se como função de perda dos modelos a *raiz quadrada do erro quadrático médio (RMSFE)* conforme Clements e Hendry (1993a, 1993b). A Figura 3 ilustra o RMSFE das previsões para os períodos sendo que as linhas tracejadas destacam os modelos univariados e as sólidas os modelos com múltiplos regressores.

Podemos observar que para os primeiros períodos de previsão os modelos estão bastante próximos, mas com vantagem dos univariados autoregressivos. Para o período 4 em diante sempre há um modelo com múltiplos regressores com o RMSFE abaixo do melhor univariado e após o 6º período nenhum multivariado apresentou maior RMSFE comparado ao melhor univariado utilizado nesse estudo. O multivariado que apresentou pior desempenho até o 6º período foi *Auto_793_FW_D_NDS* que é o único multivariado sem *dummy* de saturação. Entre o 5º e o 10º período o modelo com o menor RMSFE foi o *Auto_793_FW* o que nesse caso indica um melhor desempenho dos modelos que utilizam janela móvel.

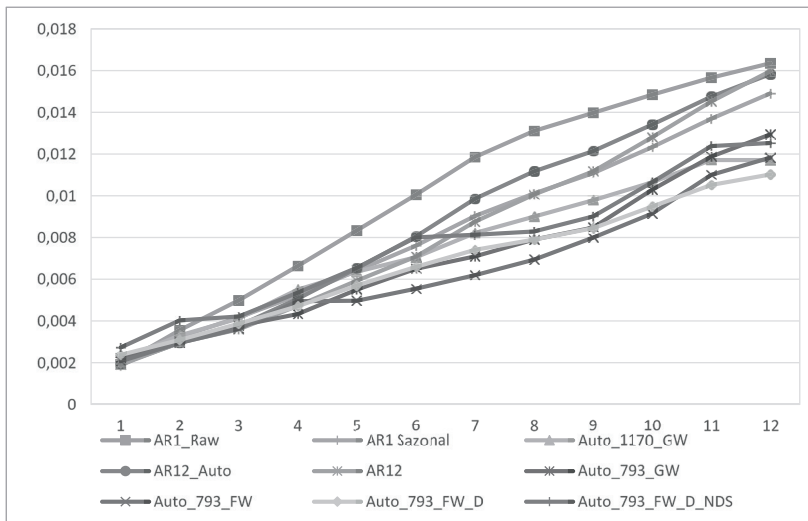


Figura 3 – RMSFE dos modelos estimados por horizonte de projeção

O modelo *Auto_793_FW_D* também chama atenção pelo seu desempenho devido à pouca variabilidade do RMSFE e que apresentou o menor erro ao final do período de 12 meses. Vale lembrar que esse modelo utiliza como variáveis explicativas uma defasagem a mais do passo a ser previsto.

Entre os modelos univariados que apresentaram menor RMSFE durante o período final de 12 meses foi o AR(1) seguido de perto pelo AR(12), ambos sem sazonalidade. O AR(1) puro nesse estudo chamado de *AR1_Raw* apresentou o pior desempenho para praticamente todo período da amostra, com melhor desempenho relativo no 12º passo.

Na Tabela 3 são listados os RMSFE referentes à Figura 3. Os modelos estão ordenados pela ordem decrescente do erro para o horizonte de projeção de 12 meses na qual é possível notar que nenhum multivariado apresentou pior desempenho que o melhor univariado

Na Tabela 4 são listados os RMSFE relativo ao *benchmark* do melhor autoregressivo descrito acima. Pode-se observar que após o período de um horizonte, na qual o AR(1) não sazonal supera todos os multivariados, pelo menos um multivariado consegue um desempenho melhor que o *benchmark*. Após seis períodos, todos os multivariados superaram o *benchmark*.

Tabela 3 – RMSFEx100 por horizonte de projeção

	Horizonte de projeção – meses											
	1	2	3	4	5	6	7	8	9	10	11	12
AR12_Auto_IIS_SIS	0.26	0.55	0.80	1.10	1.46	1.78	2.13	2.45	2.70	2.96	3.23	3.48
AR1_Raw	0.19	0.36	0.50	0.66	0.83	1.01	1.19	1.31	1.40	1.48	1.57	1.64
AR12	0.19	0.30	0.36	0.47	0.59	0.71	0.88	1.01	1.12	1.28	1.45	1.60
AR12_Auto	0.20	0.30	0.38	0.51	0.65	0.80	0.99	1.12	1.21	1.34	1.48	1.58
AR1 Sazonal	0.19	0.32	0.42	0.52	0.64	0.76	0.90	1.01	1.11	1.23	1.37	1.49
Auto_793_GW	0.22	0.30	0.38	0.43	0.55	0.65	0.71	0.79	0.85	1.03	1.19	1.29
Auto_793_FW_D_NDS	0.27	0.40	0.42	0.54	0.65	0.80	0.81	0.83	0.90	1.06	1.24	1.25
Auto_793_FW	0.21	0.30	0.36	0.50	0.50	0.55	0.62	0.69	0.80	0.91	1.10	1.18
Auto_1170_GW	0.22	0.33	0.41	0.55	0.64	0.71	0.82	0.90	0.98	1.06	1.17	1.17
Auto_793_FW_D	0.24	0.31	0.39	0.47	0.57	0.66	0.74	0.79	0.84	0.95	1.05	1.10

Fonte: Elaborado pelos autores.

Nota: Em negrito os três modelos com os menores RMSE para o período de 12 meses.

Tabela 4 – RMSFE relativo ao benchmark

	Horizonte de projeção – meses											
	1	2	3	4	5	6	7	8	9	10	11	12
Auto_793_GW	1.20	0.94	0.90	0.83	0.86	0.85	0.79	0.78	0.76	0.83	0.87	0.87
Auto_793_FW_D_NDS	1.47	1.25	1.01	1.03	1.02	1.05	0.90	0.82	0.81	0.86	0.90	0.84
Auto_793_FW	1.11	0.92	0.87	0.95	0.78	0.73	0.69	0.69	0.72	0.74	0.80	0.79
Auto_1170_GW	1.17	1.02	0.99	1.06	0.99	0.93	0.91	0.89	0.88	0.86	0.86	0.79
Auto_793_FW_D	1.23	0.87	0.78	0.71	0.68	0.66	0.62	0.60	0.60	0.64	0.67	0.67

Fonte: Elaborado pelos autores.

A menor razão entre os RMSFE foi no 8º período através do modelo Auto_793_FW_D que chegou a superar o *benchmark* em 40%, terminando o período de 12 meses com uma diferença de 33%.

Apesar do RMSFE ser uma função de perda conforme utilizado acima, ela não pode ser usada com o propósito de verificar a significância estatística dos resultados obtidos. Para tanto usaremos o critério de seleção de modelos *Model Confidence Set* que indicam o conjunto de melhores modelos (\mathcal{M}^*) dentro de um intervalo de confiança. A Tabela 5 indica o p-valor de cada modelo nos horizontes de projeção, sendo aquele considerado o melhor tem valor 1. É possível verificar que após o 5º horizonte de tempo os melhores modelos são em todos os casos os com múltiplos regressores. Nos horizontes de 10 e 12 períodos de projeção, o modelo Auto_793_FW_D consegue eliminar todos os modelos univariados com probabilidade de 80%. A Figura 4 mostra evolução dos dados realizados e previstos para alguns horizontes.

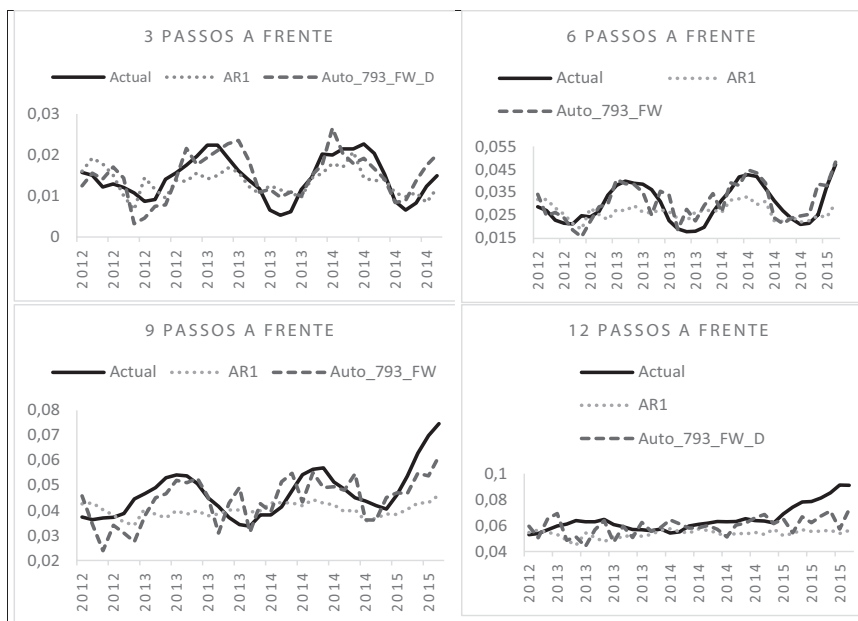


Figura 4 – Valores previstos e realizados para modelos selecionados

Tabela 5 – P-valor obtido através do Model Confidence Set

	Horizonte de projeção – meses											
	1	2	3	4	5	6	7	8	9	10	11	12
AR12_Auto_IIS_SIS	0,30	0,13	0,15	0,07	0,06	0,06	0,07	0,05	0,05	0,03	0,04	0,03
AR1_Raw	0,89	0,13	0,05	0,02	0,01	0,01	0,01	0,04	0,13	0,19	0,31	0,19
AR12	0,89	1,00	1,00	0,75	0,72	0,52	0,47	0,39	0,31	0,05	0,19	0,05
AR12_Auto	0,89	1,00	0,92	0,75	0,59	0,52	0,47	0,46	0,39	0,15	0,31	0,15
AR1 Sazonal	1,00	0,89	0,51	0,75	0,59	0,45	0,47	0,42	0,38	0,13	0,31	0,13
Auto_793_GW	0,30	1,00	0,93	1,00	0,72	0,52	0,47	0,65	0,80	0,42	1,00	0,42
Auto_793_FW_D_NDS	0,20	0,40	0,87	0,52	0,57	0,33	0,47	0,65	0,78	0,42	0,30	0,42
Auto_793_FW	0,86	1,00	0,94	0,75	1,00	1,00	1,00	1,00	1,00	0,64	0,60	0,64
Auto_1170_GW	0,43	0,89	0,87	0,67	0,57	0,33	0,08	0,07	0,26	0,64	0,60	0,64
Auto_793_FW_D	0,69	1,00	0,92	0,75	0,72	0,52	0,47	0,65	0,80	1,00	0,60	1,00

Nota 1: Os modelos pertencem ao MCS com 90% de probabilidade.

Nota 2: Em negrito, os três modelos com menores RMSFE identificados anteriormente.

Fonte: Elaborado pelos autores.

Por ser um teste estatístico, o critério de seleção de modelos MCS é sensível ao tamanho da amostra, podendo esse ser um dos motivos pelo qual o resultado desse teste não seja mais expressivo em horizonte mais longo em linha com o RMSFE. Também é possível perceber que para maiores horizontes de tempo, os modelos com múltiplos regressores são melhores ranqueados que os autoregressivos no conjunto final.

O número médio de variáveis e *dummies* presentes nos modelos terminais finais estão listados na Tabela 6. Nota-se que estes números não variam muito com o aumento do horizonte de projeção, o que faz sentido considerando que esses números são do modelo Auto_793_FW_D com janela móvel, ou seja, amostra de tamanho fixo. O menor número médio de variáveis foi para o horizonte 5 com 7,64 variáveis e no maior caso 8,89 para o 11º período. A quantidade de *dummies* também estão entre 1,55 e 2,41.

Tabela 6 – Número médio de variáveis e *dummies* nos modelos finais

	Horizonte de projeção – meses											$\alpha=0,0001$
	1	2	3	4	5	6	7	8	9	10	11	
Variáveis	7.54	8.11	7.82	7.77	7.64	6.80	6.85	8.62	7.29	8.46	8.89	6.77
Dummies	2.04	2.17	1.96	1.90	1.55	2.10	1.56	1.70	1.67	2.41	1.75	2.16

Fonte: Elaborado pelos autores.

Com relação as 793 variáveis do modelo Auto_793_FW_D, verificou-se que 346 delas foram, pelo menos uma vez, retidas nos modelos terminais finais.

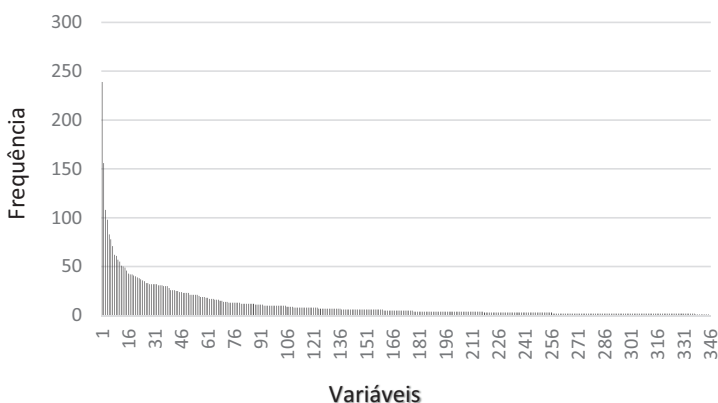


Figura 5 – Distribuição de frequência das variáveis do modelo Auto_793_FW_D

A Figura 5 ilustra a distribuição de frequência das variáveis que foram selecionadas. Pode-se observar que algumas variáveis foram retidas inúmeras vezes, 239 no maior caso, e outras foram selecionadas apenas uma ou duas vezes.

Apesar do presente trabalho não ter como objetivo testar teorias econômicas, mas tendo em vista explicar o motivo pelo qual algumas foram retidas em detrimento de outras, foram listadas na Tabela 7 as vinte variáveis com maior incidência de seleção. Do lado fiscal algumas variáveis como 4716, 5456, 5761, 5434 que se referem à necessidade de financiamento do setor público do governo federal, banco central, estados e municípios tanto internos quanto externos, aparecem de forma recorrente. Outras variáveis têm relação com atividade econômica e consumo como a 4382. Do lado monetário temos 1830, 1829, 1848.

Tabela 7 – Variáveis com maior frequência no modelo Auto_793_FW_D

CÓDIGO	DESCRIÇÃO	UNIDADE	Freq
2226	Dívida mobiliária (Saldos) - Títulos do Tesouro Nacional - Posição de custódia - CDP	u.m.c. (milhões)	239
4218	Exportações (Kg) - Soja mesmo triturada	Kg	156
10702	Dívida mobiliária federal (saldos) - Posição de custódia - Título indexado a TR - NTN-H	u.m.c. (milhões)	108
1848	Recolhimentos obrigatórios de instituições financeiras - Depósitos de poupança em espécie (remunerados)	u.m.c. (mil)	98
7409	Inadimplência - Taxa líquida (t-3) = (novos-cancelados) /consultas	%	83
4382	PIB acumulado dos últimos 12 meses - Valores correntes (R\$ milhões)	R\$ (milhões)	78
1830	Base monetária ampliada - Títulos do Banco Central-financiamento líquido (saldo em final de período)	u.m.c. (mil)	71
1829	Base monetária ampliada - Títulos do Banco Central-carreira do mercado (saldo em final de período)	u.m.c. (mil)	62
13072	Saldo da balança comercial - Distrito Federal	US\$ (mil)	61
2971	Exportações - Soja mesmo triturada	US\$	57
4392	CDI acumulada no mês anualizada base 252	% a.a.	55
4716	NFSP - Necessidade de Financiamento Setor Público sem desvalorização cambial - Fluxo acumulado no ano - Resultado nominal - Total - Governo Federal e Banco Central	R\$ (milhões)	51
13522	Índice nacional de preços ao consumidor - amplo (IPCA) - em 12 meses	%	50
5456	NFSP - Necessidade de Financiamento Setor Público sem desvalorização cambial (% PIB) - Fluxo acumulado no ano - Resultado nominal - Externa - Governos estaduais e municipais	%	49
5761	NFSP - Necessidade de Financiamento Setor Público sem desvalorização cambial (% PIB) - Fluxo acumulado em 12 meses - Juros nominais - Interna - Governo Federal e Banco Central	%	46
11771	Índice da taxa de câmbio real (INPC) - Jun./1994=100 - Peso argentino - mensal	Índice	43
1838	Meios de pagamento amplos - Quotas de fundos de renda fixa (saldo em final de período)	u.m.c. (mil)	42
5434	NFSP - Necessidade de Financiamento Setor Público sem desvalorização cambial (% PIB) - Fluxo acumulado no ano - Resultado nominal - Total - Governos estaduais e municipais	%	42
4195	Exportações (Kg) - Algodão em bruto	Kg	41

Fonte: Elaborado pelos autores.

A Tabela 8 enumera todas as *dummies* que foram retidas nos modelos finais ordenadas por frequência de ocorrência. É possível observar que no período do segundo semestre de 2002 ao primeiro de 2003, temos uma grande concentração das *dummies* que indicam as quebras. Isso pode ser explicado dado o fato que nesse período houve um choque cambial no Brasil, tendo os juros elevados entre outubro de 2002 a fevereiro de 2003 de 18% a 26,5% e a inflação em maio de 2003 chegou a 16,5%.

Tabela 8 – Variáveis com maior frequência no modelo

Dummy	Freq	Dummy	Freq	Dummy	Freq	Dummy	Freq
S1:2002(10)	53	S1:2006(4)	3	S1:2005(11)	1	S1:2011(10)	1
I:2002(11)	52	S1:2011(6)	3	I:2004(7)	1	S1:2007(10)	1
S1:2003(1)	39	S1:2014(12)	3	I:2003(4)	1	S1:2008(6)	1
S1:2006(2)	35	S1:2013(6)	3	S1:2004(4)	1	I:2002(10)	1
S1:2003(4)	30	S1:2010(9)	3	S1:2004(7)	1	S1:2005(6)	1
S1:2003(5)	28	S1:2011(4)	3	S1:2011(11)	1	I:2004(12)	1
S1:2003(3)	24	S1:2015(3)	2	I:2003(3)	1	S1:2010(10)	1
S1:2002(9)	18	S1:2010(4)	2	S1:2008(2)	1	S1:2006(12)	1
S1:2010(8)	9	S1:2002(12)	2	S1:2013(4)	1	S1:2002(8)	1
S1:2005(5)	9	I:2010(6)	2	S1:2012(7)	1	S1:2007(5)	1
S1:2002(11)	8	I:2008(6)	2	S1:2008(8)	1	S1:2003(6)	1
S1:2005(4)	7	S1:2012(2)	2	S1:2009(4)	1	S1:2006(6)	1
S1:2003(2)	7	S1:2015(1)	2	S1:2006(8)	1	S1:2011(3)	1
I:2003(1)	6	I:2005(10)	2	S1:2012(12)	1	S1:2014(3)	1
S1:2010(5)	4	I:2004(6)	2	S1:2006(5)	1		
S1:2005(1)	4	I:2005(4)	2	S1:2014(9)	1		
I:2002(12)	4	S1:2006(1)	2	S1:2013(1)	1		
S1:2005(2)	4	I:2003(6)	2	S1:2012(5)	1		
S1:2011(5)	4	S1:2010(7)	1	S1:2013(5)	1		
S1:2006(3)	3	S1:2004(5)	1	S1:2012(10)	1		

Fonte: Elaborado pelos autores.

6. Possíveis Extensões e Limitações do Trabalho

O presente trabalho procurou explorar a grande quantidade de dados disponíveis atualmente para efetuar a previsão de inflação em modelo com regressores múltiplos e verificar se essas informações adicionais poderiam

apresentar ganhos se comparados a modelos univariados. Também houve a preocupação com relação à comparação de modelos evitando fazê-lo somente com funções de perda apenas, mas utilizando o MCS.

Uma das limitações encontradas no estudo é referente ao relativo curto período de tempo das previsões que se iniciam em novembro de 2011 até setembro de 2015 totalizando 36 previsões de um a doze passos à frente. A principal restrição nesse sentido é com relação à grande parte das séries de tempo utilizadas não estarem disponíveis em períodos históricos mais longos. Conforme pode ser observado na Figura 2, o número de dados disponíveis pelo Banco Central cresce com o passar do tempo e grande parte das utilizadas nesse estudo não estão disponíveis em períodos anteriores a década de 90, por exemplo. Nesse sentido, o MCS teve um papel fundamental para indicar quão informativas são as séries utilizadas definindo a significância dos resultados dado o tamanho da amostra. Outro ponto a ser observado é que as previsões não contemplam períodos de grande turbulência como a crise de 2008 ou a alta no preço das commodities em 2010, sendo que a situação observada entre 2011 e 2015 é acerca de um período de alta na inflação ultrapassando de forma consistente o centro da meta preestabelecida pelo Banco Central do Brasil

Como extensão do trabalho pode-se citar recentes avanços em previsões de inflação utilizando dados desagregados como em Carlo e Marçal (2016), que obtiveram evidências de ganhos de previsão utilizando as séries desagregadas do IPCA. O mesmo procedimento desse trabalho poderia ser aplicado às séries desagregadas do IPCA para verificar se as inúmeras variáveis econômicas possibilitariam ganhos de previsão sobre a série agregada, já que em Carlos e Marçal (2016) os modelos utilizados foram todos univariados.

Outra extensão possível é dada por modelos FAVAR como em Figueiredo e Guillén (2013). Os autores aplicaram a metodologia para prever a inflação brasileira o período de 2008 e 2012 com 40, 64 e 415 variáveis econômicas selecionadas a partir de testes de causalidade de Granger. Os resultados deste trabalho são, em certa medida, semelhantes aos encontrados Figueiredo e Guillén (2013). Os autores encontraram ganhos de RMFSE para prazos principalmente posteriores a 12 meses, mas não utilizaram o MCS.

Outro estudo que utilizou o *Autometrics* mas para prever a inflação dos Estados Unidos entre 1959 e 2006, com dados trimestrais, utilizando variáveis, fatores, ambos e modelos autoregressivos univariados, pode ser encontrada em Castle, Clements e Hendry (2011). Nesse trabalho, os autores concluíram que o autoregressivo de ordem 1 foi o que gerou o menor RMSFE. Apesar do aparente melhor desempenho do autoregressivo, nesse trabalho também não foi utilizado o MCS estando sujeito a mesma crítica que Hansen, Lunde e Nason (2010) fizeram a Stock e Watson (1999) que talvez a série não seja informativa o suficiente para diferenciar os métodos aplicados.

7. Conclusão

Neste trabalho, foi utilizado o algoritmo *Autometrics* para selecionar entre mais de mil variáveis econômicas as que são mais relevantes dentro da amostra para explicar o IPCA e utilizá-las para previsão fora da amostra para períodos de um a doze meses à frente. A fim de comparar o desempenho dos modelos e avaliar o ganho de desempenho ao utilizar um grande conjunto de variáveis, o IPCA também foi projetado utilizando técnicas clássicas de previsão de séries de tempo como autoregressivos de ordem 1, 12 com e sem sazonalidade, estimados e não pelo *Autometrics*, com e sem *dummy* de saturação.

A comparação dos modelos foi determinada pelo *Model Confidence Set* desenvolvido por Hansen *et al.* (2010), que indicam ganhos de projeções utilizando variáveis sendo elas mais expressivas para horizontes mais longos no tempo. O modelo que contém a mais ampla gama de regressores tem bom desempenho para todos os horizontes, principalmente para horizonte mais longos.

Este trabalho procura contribuir com a literatura de previsão de inflação brasileira. A utilização com método de um grande conjunto de dados para esse caso mostra-se benéfica e sugere a existência de ganhos de previsão para horizontes frente a *benchmarks* ingênuos ou modelos univariados.

Referências

- Akaike, H. 1973. "Information theory and an extension of the likelihood principle". In B. Petrov and F. Czaki (Eds.), *Second International Symposium on Information Theory*, pp. 267-287. Budapest: Akademia Kiado.
- Atkeson, A., & Ohanian, L. E. 2001. "Are Phillips curves useful for forecasting inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review* 25(1): 2-11.
- Bader, F. L. C., Koyama, S. M., & Tsuchida, M. H. 2014. "Modelo FAVAR Canônico para Previsão do Mercado de Crédito" 369.
- Bernanke, B., Boivin, J., Elias, P. 2005. "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach". *Quarterly Journal of Economics* 120: 387-422.
- Box, G. E. P., Jenkins, G. M. 1976. "Time Series Analysis: Forecasting and Control" (2nd ed.). San Francisco, CA: Holden-Day.
- Campos, J., & Ericsson, N. R. 1999. "Constructive data mining: modeling consumers' expenditure in Venezuela". *The Econometrics Journal* 2(2): 226-240.
- Campos, J., Hendry, D. F., & Krolzig, H. M. 2003. "Consistent Model Selection by an Automatic Gets Approach". *Oxford Bulletin of Economics and Statistics* 65(s1): 803-819.
- Carlo, T. C.; Marçal, E. F. 2016. "Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon". *Applied Economics* 48(50): 4846-4860.
- Castle, J.L., Hendry, D.F., Clements, M. P. 2011. "Forecasting by Factors, by Variables, by Both, or Neither?" Working paper, Economics Department, University of Oxford.
- Castle, J. L., Qin, X., & Robert Reed, W. 2013. "Using model selection algorithms to obtain reliable coefficient estimates". *Journal of Economic Surveys* 27(2): 269-296.
- Castle, J. L., Doornik, J. A., & Hendry, D. F. 2011. "Evaluating automatic model selection". *Journal of Time Series Econometrics* 3(1).
- Cecchetti, S. G., Chu, R. S., & Steindel, C. 2000. "The unreliability of inflation indicators". *Current Issues in Economics and Finance* 6(4): 1-6.
- Clements, M. P., & Hendry, D. F. 1993. "On the limitations of comparing mean square forecast errors". *Journal of Forecasting* 12(8): 617-637.
- Dees, S., Mauro, F. D., Pesaran, M. H., & Smith, L. V. 2007. "Exploring the international linkages of the euro area: a global VAR analysis". *Journal of Applied Econometrics* 22(1): 1-38.
- Dickey, D. A., & Fuller, W. A. 1979. "Distribution of the estimators for autoregressive time series with a unit root". *Journal of the American statistical association* 74(366a): 427-431.
- Doornik, J. A., & Ooms, M. 2007. "Introduction to Ox: An Object-Oriented Matrix Language".
- Doornik, J. A. 2009. "Autometrics. In in Honour of David F. Hendry".
- Doornik & Hendry. 2014. "Empirical Model Discovery and Theory Evaluation, Arne Ryde Memorial Lectures", MIT Press.
- Doornik & Hendry. 2015. "Statistical model selection with "Big Data" *Cogent Economics & Finance* 3: 1045216
- Doornik, J. A., Hendry, D. F., & Pretis, F. 2013. "Step-indicator saturation". Discussion Paper 658.
- Figueiredo, F.M.R. 2010. "Forecasting Brazilian inflation using a large data set". Central Bank of Brazil Working Paper 228.
- Figueiredo, F.M.R., Guillén, O.T.C. 2013. "Forecasting Brazilian consumer inflation with FAVAR models using target variables", mimeo.
- Granger, C. W. J., Anderson, A. P. 1978. "Introduction to Bilinear Time Series Models". Vandenhoeck and Ruprecht.

- Hansen, P. R. 2005. "A Test for Superior Predictive Ability". *Journal of Business and Economic Statistics* 23: 365–380.
- Hansen, P. R., Lunde, A., & Nason, J. M. 2010. "The model confidence set". Available at SSRN 522382.
- Hendry, D. F. 1987. "Econometric methodology: A personal perspective". *Advances in econometrics* 2: 29-48.
- Hendry, D. F. 1995. "Dynamic econometrics". Oxford University Press.
- Hendry, D. F. 2005. "Predictive failure and econometric modelling in macroeconomics: The transactions demand for money". In *General-to-specific modelling*, pp. 535-560, Edward Elgar.
- Hendry, D. F., & Nielsen, B. 2007. "Econometric modeling: a likelihood approach". Princeton University Press.
- Hoover, K. D., & Perez, S. J. 1999. "Data mining reconsidered encompassing and the general-to-specific approach to specification search". *The Econometrics Journal* 2(2): 167-191.
- Hendry, D. F., & Krolzig, H. M. 1999. "Improving on Data mining reconsidered by KD Hoover and SJ Perez". *The Econometrics Journal* 202-219.
- Kozicki, S., & Tinsley, P. A. 2001. "Shifting endpoints in the term structure of interest rates". *Journal of Monetary Economics* 47(3): 613-652.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., & Winkler, R. 1982. "The accuracy of extrapolation (time series) methods: Results of a forecasting competition". *Journal of Forecasting* 1(2): 111-153.
- Makridakis, S., & Hibon, M. 2000. "The M3-Competition: results, conclusions and implications". *International Journal of Forecasting* 16(4): 451-476.
- Romano, J. P., & Wolf, M. 2005. "Stepwise multiple testing as formalized data snooping". *Econometrica* 73(4): 1237-1282.
- Pesaran, M. H., Schuermann, T., & Smith, L. V. 2009. "Forecasting economic and financial variables with global VARs". *International Journal of Forecasting* 25(4): 642-675.
- Santos, C., Hendry, D. F., & Johansen, S. 2008. "Automatic selection of indicators in a fully saturated regression". *Computational Statistics* 23(2): 317-335.
- Sims, C. A. 1980. "Macroeconomics and Reality". *Econometrica: Journal of the Econometric Society* 1-48.
- Stock, J. H., & Watson, M. W. 1989. "New indexes of coincident and leading economic indicators". In *NBER Macroeconomics Annual 1989(4)*: 351-409. MIT press.
- Stock, J. H., Watson, M. W. 1998. "Diffusion indexes". Working Paper 6702, NBER.
- Stock, J. H., Watson, M. W. 2002. "Macroeconomic forecasting using diffusion indices". *Journal of Business and Economic Statistics* 20(2): 147–162.
- Stock, J. H., Watson, M.W. 1999. "Forecasting inflation". *Journal of Monetary Economics* 44: 293–335.
- Stock, J. H., & Watson, M. W. 2006. "Forecasting with many predictors". *Handbook of Economic Forecasting* 1: 515-554.
- Stock, J. H., & Watson, M. 2009. "Forecasting in dynamic factor models subject to structural instability". *The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry* 173, 205.
- White, H. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". *Econometrica: Journal of the Econometric Society* 817-838.
- White, H. 2000. "A Reality Check for Data Snooping". *Econometrica* 68: 1097– 1126.
- White, H. 1990. "A consistent model selection procedure based on m-testing". *Modelling Economic Series: Readings in Econometric Methodology*, Clarendon Press, Oxford, 369-383.