

FUNDAÇÃO GETÚLIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

DOUGLAS MINORU KAGOHARA

**AVALIANDO TÉCNICAS DE NOWCASTING: UMA APLICAÇÃO DO PIB
BRASILEIRO**

SÃO PAULO

2015

DOUGLAS MINORU KAGOHARA

**AVALIANDO TÉCNICAS DE NOWCASTING: UMA APLICAÇÃO DO PIB
BRASILEIRO**

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

Linha de pesquisa: Macroeconomia aplicada

Orientador: Prof. Dr. Emerson F. Marçal

SÃO PAULO

2015

Kagohara, Douglas Minoru.

Avaliando técnicas de nowcasting: Uma aplicação do PIB brasileiro / Douglas Minoru Kagohara. - 2015.

48 f.

Orientador: Emerson Fernandes Marçal

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Produto interno Bruto - Brasil. 2. Previsão. 3. Macroeconomia. 4. Modelos não-lineares (Estatísticas). I. Marçal, Emerson. II. Dissertação (MPFE) - Escola de Economia de São Paulo. III. Título.

CDU 330.55(81)

DOUGLAS MINORU KAGOHARA

**AVALIANDO TÉCNICAS DE NOWCASTING: UMA APLICAÇÃO DO PIB
BRASILEIRO**

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

Linha de pesquisa: Macroeconomia aplicada

Orientador: Prof. Dr. Emerson F. Marçal

Data da aprovação: 13/08/2015

Banca examinadora:

Prof. Dr. Emerson F. Marçal (orientador)
FGV - EESP

Prof. Dr. Frederico Turolla
ESPM - Departamento de Economia

Prof. Dr. Pedro Luiz Valls Pereira
FGV - EESP

AGRADECIMENTOS

Gostaria de agradecer a todos que contribuíram de alguma maneira para a conclusão deste trabalho e do curso de mestrado, e fazem parte da minha vida em maior ou menor intensidade.

Em primeiro lugar, agradeço à pessoa que escolhi para estar ao meu lado, minha esposa Maria Fernanda Namour, por todo o apoio oferecido antes e durante o curso, pela compreensão da minha ausência em alguns períodos e por ser meu suporte nos períodos de dificuldade. Estes fatores foram decisivos para a conclusão do curso e aproveitamento de todas as oportunidades que este ofereceu.

Além disto, agradeço aos meus pais Milton e Cida pelo apoio incondicional à continuidade dos estudos e aperfeiçoamento constante do conhecimento.

Agradeço também a todos os colegas de turma e professores, pela troca intensa de experiências, dificuldades e conhecimentos. Ajudaram muito na motivação e tornaram o convívio mais gratificante ao longo do curso.

Um agradecimento especial ao meu orientador Prof. Emerson Marçal, pela disponibilidade e ajuda desde o início do trabalho, oferecendo apoio constante ao longo do mesmo. Com certeza, seus conselhos foram de extrema importância no direcionamento do trabalho e análise dos resultados.

Por último, gostaria de agradecer aos professores que estiveram presentes na banca de avaliação, prof. Pedro Valls e Frederico Turolla. Seus comentários foram extremamente relevantes, e permitiram aperfeiçoar ainda mais o trabalho.

Espero ter retribuído a todos da mesma maneira, e na mesma intensidade.

RESUMO

O objetivo deste trabalho é aplicar e avaliar o desempenho do conceito de técnicas de *nowcasting* para previsão de uma importante variável macroeconômica do Produto Interno Bruto (PIB) brasileiro. Nos últimos anos, novas técnicas vêm sendo propostas e aprimoradas. Comparam-se diferentes modelos de *nowcasting* frente a um *benchmarking*, avaliando a relevância das variáveis a partir do *Autometrics*, que foi desenvolvido por Doornik (2011).

A proposta é reunir diversos indicadores econômicos da economia brasileira que possam em maior ou menor grau antecipar a variação do PIB. Será utilizada a técnica de variáveis *dummies* com saturação (proposta por Johansen et. al.) para controlar possíveis quebras e *outliers*. Esta abordagem é adequada para um ambiente econômico instável, com constantes mudanças ao longo do tempo.

Palavras-chave: Produto Interno Bruto (PIB) do Brasil; *Nowcasting*; *Forecasting*; *Autometrics*; variáveis *dummies* com saturação; *Bridge Equations*;

ABSTRACT

The aim of this work is to apply and evaluate the performance of nowcasting techniques for prediction of an important macroeconomic variable of gross domestic product (GDP). In recent years, new techniques have been proposed and improved. We will compare different models of nowcasting against a benchmark, evaluating the variables are relevant from Autometrics, which was developed by Doornik (2011).

The proposal is to bring together various economic indicators of the Brazilian economy that might to a greater or lesser extent anticipate GDP growth. It will be used the technique of dummy variables with saturation (proposed by Johansen et. Al.) to control possible breaks and outliers. This approach is suitable under an unstable economic environment with constant change over time.

Key-Words: Brazilian Gross Domestic Product (GDP); Nowcasting; Forecasting; Autometrics; *dummies* with saturation; Bridge Equations;

Lista de Figuras

1	Representações gráficas da série do PIB trimestral do Brasil - fonte: Ipeadata .	26
2	Exemplo de defasagens de publicação dos indicadores - IBOVESPA e produção industrial.	28
3	Gráfico mostrando a estacionariedade de algumas das variáveis utilizadas nos modelos deste trabalho	29
4	Comparação do EQM para os modelos.	32
5	Gráfico com os erros de previsão por modelo.	33
6	Testes de Diebold-Mariano	34
7	Teste de Diebold-Mariano para comparação com o modelo utilizado como <i>benchmarking</i>	36
8	Lista completa dos indicadores utilizados neste trabalho, categorizadas de acordo com seu Tipo, Grupo, Periodicidade, início da série, período de divulgação dentro do mês, defasagem da informação e órgão responsável.	44
9	Defasagens das variáveis explicativas utilizadas nos modelos de <i>forecasting</i> e <i>nowcasting</i>	45

Lista de Tabelas

1	Indicadores mensais utilizados como variáveis explicativas nos modelos	27
2	Modelos que serão utilizados neste trabalho, com a especificação de suas respectivas variáveis explicativas.	30
3	Resultado do EQM dos 6 modelos por ciclo econômico.	36
4	EQM do modelo NC3 por ciclo econômico.	37
5	Principais variáveis no modelo final com seus respectivos coeficientes, defasagens e significância.	39
6	Principais variáveis no modelo final com seus respectivos coeficientes, defasagens e significância.	41

Sumário

1	Introdução	10
2	Revisão da Literatura	12
3	Referencial Teórico e Especificação dos Modelos	16
3.1	Nowcast	16
3.2	Modelagem Matemática	18
4	Seleção automática de modelos	20
4.1	<i>Autometrics</i>	20
4.2	Impulse-Indicator Saturation (IIS)	21
4.3	Teste de Diebold-Mariano (DM) para comparar previsões	22
5	Metodologia	25
5.1	Base de Dados	25
5.2	Especificação dos modelos	29
6	Resultados	32
6.1	Modelos	32
6.2	Ciclos Econômicos	36
6.3	Variáveis finais do modelo	37
6.4	Sugestão para pesquisas	42
7	Conclusão	43
A	Apêndice	44
A.1	Lista de indicadores e suas defasagens utilizados no modelo	44
A.2	Ciclos Econômicos	45
B	Bibliografia	47

1 Introdução

Nowcasting é um termo em inglês derivado da junção entre as palavras *Now* (agora) e *Forecasting* (previsão). Esta abordagem é relativamente recente em aplicações na área econômica, e consiste em prever o passado, presente e futuro de curto prazo. Em geral, quando falamos de períodos econômicos de certa estabilidade, as técnicas de *forecasting* têm bom desempenho preditivo. Porém, para períodos com mudanças mais frequentes, estas previsões podem se deteriorar com o passar do tempo.

A proposta do conceito de *nowcasting* é aproveitar a disponibilidade de indicadores de frequência mais alta para atualização da previsão de indicadores econômicos com menor frequência, como o Produto Interno Bruto (PIB) de um país, por exemplo, que em geral, tem divulgação trimestral e com defasagens em relação ao período considerado. Dada a alta velocidade e demanda por informação nos mercados financeiros reais, esperar quase 5 meses (no caso do Brasil) para a publicação do PIB Real do trimestre anterior, pode não ser viável para algumas aplicações - decisão de política econômica, por exemplo. Neste sentido, pode-se explorar o conteúdo preditivo de indicadores e variáveis com atualização mais frequente (mensal, diária) como, por exemplo, o índice de ações, produção industrial e pesquisas de confiança, que são boas *proxies* para a previsão do PIB, para obter estimativas relativamente precisas e estáveis desta ou de outras importantes variáveis macroeconômicas ¹, e o mais importante: em períodos de tempo mais frequente que o trimestral, e que pode ser constantemente atualizado com a publicação de novos *releases* de informação.

O objetivo deste trabalho é realizar um exercício de *nowcasting* para previsão do PIB Real brasileiro, e está dividido da seguinte maneira. Além desta introdução, no capítulo 2 faremos uma revisão da literatura sobre modelos e técnicas de *nowcasting*, e a seção 3 apresenta a teoria que suporta todo o arcabouço estatístico para *nowcast* e suas aplicações. Na seção 4 iremos discutir o conceito do algoritmo *Autometrics*, e o Impulse Indicator Saturation (IIS) - técnica utilizada para tratamento de quebras estruturais nas séries - além de apresentar o teste formal de Diebold-Mariano para comparação de previsão entre modelos. A seção 5 descreve a metodologia e a base de dados que será utilizada neste trabalho, incluindo a formulação matemática

¹Mercado de trabalho, inflação, dívida pública, etc.

do nosso problema, enquanto que a seção 6 discute os resultados obtidos. Finalmente, a seção 7 apresenta as principais conclusões deste trabalho, e os materiais de apoio encontram-se no Apêndice da seção A.

2 Revisão da Literatura

A utilização do conceito de *nowcasting* vem sendo apresentado na literatura há algum tempo para previsão da variação de fatores climáticos, e na área de economia apresenta um uso relativamente recente na previsão de variáveis macroeconômicas, tanto no contexto do Brasil quanto em outros países. Apesar das vantagens na aplicação desta técnica, temos diversos desafios a enfrentar e controlar, que são largamente discutidos em (Castle, Hendry e Kitov, (2013)) e (Banbura, Giannone e Modugno, (2013)). Em geral, o número de variáveis envolvidas é muito grande, e com diferentes frequências de publicação. Além disto, como estamos tratando de variáveis com diversas origens e não necessariamente sincronizadas, a última informação disponível pode não ser contemporânea, o que causa o problema chamado de "ragged edge" na literatura. Estes são alguns dos exemplos que devem ser contornados para termos bons modelos de *nowcasting*.

O trabalho de (Castle, Hendry e Kitov, (2013)) discute exaustivamente alguns dos principais desafios para desenvolvimento de bons modelos de previsão de variáveis macroeconômicas que são:

1. "*missing data problem*", ou a falta de informação de alguns indicadores para previsão da variável em questão;
2. "*latency problem*", ou a diferença de tempo entre a publicação dos diversos indicadores;
3. "*changing database problem*", que diz respeito à constante atualização ou correção dos dados já publicados;
4. "*measurement error problems*" ou a falta de acurácia na obtenção dos dados preliminares;
5. "*Breaks Problem*" ou a quebra estrutural - mudanças nas médias das variáveis analisadas;
6. "*Problem of more variables, N, than observations*" no caso de variáveis de alta frequência e
7. "*Non-synchronous release problem*" ou a dificuldade em lidar com diferentes defasagens das variáveis utilizadas

Neste artigo, os autores propõem tratamentos para cada um destes pontos chave, de maneira a aliar boas previsões com o conceito de *nowcasting*. A interação entre todos estes pontos também é uma dificuldade adicional desta técnica. Para comprovação dos modelos e tratamentos apresentados, os autores apresentam resultados empíricos que utilizaram a técnica de *nowcasting* para a previsão do crescimento do PIB do Reino Unido, endereçando cada um dos pontos de dificuldades listados acima.

Neste sentido, (Liu, Matherson e Romeu, (2011)) também testaram 5 modelos de *forecasting* e *nowcasting* do PIB de 10 países da América Latina (representando 94% do total da região), mostrando a importância e relevância do uso das informações de indicadores mensais para aumentar a eficácia destas previsões. Como resultado geral, o *nowcast* mostrou-se mais assertivo nos modelos cujas informações dos indicadores estivessem mais completas, ou seja, o erro de previsão para o terceiro mês é menor que o primeiro se considerarmos o período de um trimestre. Outro resultado importante foi o destaque para a importância da utilização de variáveis explicativas externas e de preços de commodities, indicando um link relevante destas economias com os EUA. Isto reforça a importância da utilização de tal tipo de informação em nosso estudo. Também temos trabalhos que aplicam as técnicas de *nowcasting* para previsão de variáveis macroeconômicas em outros países: (i) PIB da China em (Yiu e Chow, (2011)) - no qual os autores trabalham com 189 séries de indicadores econômicos chineses (ii) PIB da Noruega em (Aastveit e Trovik, (2007))- economia pequena, aberta e forte exportadora de petróleo, no qual os autores mostraram que o índice de ações neste país é o conjunto de dados mais importantes para atualização da previsão do PIB trimestral (indicadores de mercado de trabalho e produção industrial também foram significativos), e recentemente (iii) do PIB da Turquia em (Modugno, Soybilgenx e Yazgan, (2015)), no qual os autores estimam o PIB trimestral e anual, considerando 14 dos principais indicadores econômicos daquele país.

No caso do Brasil, o trabalho de (Bragoli, Metelli e Modugno, (2014)) avaliou o desempenho do indicador de pesquisa diária do PIB (realizado pelo Banco Central) comparada à modelos clássicos de *nowcasting* que apresentaram bons resultados na literatura. Também mostraram a importância de utilizar informações e variáveis de maior frequência de publicação (questionários, pesquisas de confiança, produção industrial, etc) do que o próprio PIB para aumentar a

eficácia das previsões, e aproveitando-se de informações valiosas antes da publicação da própria informação do PIB. Os autores avaliaram a pesquisa diária do PIB e outros indicadores como preditivos do PIB Real no próximo período dentro da ótica de *nowcast*: estimar a variável de interesse no intervalo de tempo entre o início do trimestre de referência e a sua publicação por órgãos oficiais do governo, explorando informações de indicadores e outras variáveis de maior frequência. O resultado mostrou que tais informações são estatisticamente tão boas quanto *forecasts* baseados em modelos tradicionais. No entanto, os autores se limitaram a testar somente algumas variáveis determinadas como relevantes, e outra parte originárias de órgãos internacionais de divulgação, como a Bloomberg, por exemplo.

Outros trabalhos na literatura têm foco na previsão de diferentes variáveis macroeconômicas, como a inflação anual na zona do Euro em (Banbura, Giannone e Reichlin, (2010)) e americana em (Giannone, Reichlin e Small, (2006)), uma aplicação para previsão de ciclos econômicos em (Em progresso - D'Agostino, Giannone, Lenza e Modugno, (2014)) e previsão de crescimento econômico global - estimado pelo Fundo Monetário Internacional (FMI) - em (Ferrara e Marsilli, (2014)), no qual os autores trabalharam com uma grande base de indicadores mensais (392 ao todo, incluindo indicadores financeiros, reais da economia, pesquisas de confiança e gerais) de vários setores para a economia de 37 países, representando mais de 80% do PIB Mundial.

Considerando a literatura de modelos de *forecasting* tradicional, (Ericsson, (2012)) analisou técnicas de *forecasting* da dívida do governo americano, apresentando os riscos e impactos dos erros de previsão, e a importância da acurácia de previsão deste importante indicador macroeconômico. Segundo o autor, erros de previsão são mais comuns nos pontos de inversão no ciclo de negócios e expansões econômicas, e defende a importância da técnica chamada de *IIS - impulse indicator saturation* na detecção de quebras estruturais nas variáveis. Esta técnica é um teste genérico para um número desconhecido de quebras, que podem ocorrer em tempos, duração e magnitudes desconhecidas, e a qualquer momento na amostra. Estes resultados motivaram a utilização desta técnica neste trabalho.

Em geral, temos na literatura evidências do uso de técnicas de *nowcasting* com diversas conclusões, conforme discutido em (Banbura, Giannone e Modugno, (2013)). Os resultados

principais mostram que este tipo de modelagem tem melhor desempenho quando consideramos um horizonte relativamente curto de tempo, em particular do período atual, e o procedimento estatístico é tão consistente quanto os modelos de *forecasting* tradicionais. Além disto, a assertividade da previsão aumenta consideravelmente ao final dos períodos, indicando a importância de atualização das informações tão logo estejam disponíveis.

3 Referencial Teórico e Especificação dos Modelos

3.1 Nowcast

Para formulação e resolução dos modelos de *nowcasting*, a literatura em geral tem uma abordagem tradicional que é a solução de “*bridge equations*”, também bastante utilizada junto à análise julgamental de diversos bancos centrais segundo (Banbura, Giannone e Reichlin, (2010)). Este é um arcabouço bem simplificado de equações, no qual o *nowcast* é obtido através de uma regressão da variável objetivo trimestral em relação às suas respectivas defasagens e alguns indicadores mensais. Neste caso, teremos apenas informações parciais mensais para um dado trimestre, e é necessária alguma técnica de interpolação para preenchimento dos dados faltantes. Alguns trabalhos como (Ferrara, Guegan e Rakotomalahy, (2010)) têm uma abordagem mais elaborada para este tipo de modelo, desenvolvendo um modelo que se mostra robusto para *nowcasts* e *forecasts* da taxa de crescimento trimestral do PIB, através de modelagem não paramétrica, baseados na vizinhança próxima para prever as variáveis mensais envolvidas na modelagem paramétrica do PIB utilizando “*bridge equations*”.

No entanto, podemos seguir uma abordagem mais sofisticada como em (Castle, Hendry e Kitov, (2013)) e (Banbura, Giannone e Reichlin, (2010)). Vamos ilustrar e discutir o problema de modelar o *nowcasting* de PIB como exemplo. Esta variável é de publicação trimestral na maioria dos países, e o que queremos é estimá-la com base em indicadores de maior frequência dentro do período considerado. Formalmente: Ω_v é o conjunto de informações disponíveis no período v . Denotando o crescimento do PIB no período t por y_t^Q . O problema de *nowcasting* de y_t^Q é a projeção ortogonal de y_t^Q no conjunto de informações disponíveis Ω_v , de tal maneira que:

$$P \left[y_t^Q | \Omega_v \right] = E \left[y_t^Q | \Omega_v \right] \quad (1)$$

e $E[\cdot | \Omega_v]$ é a esperança condicional. Relembrando, o universo Ω_v é um conjunto de dados que podem ser assíncronos e publicados em diferentes momentos, e pode conter séries com diferentes periodicidades. Logo teremos:

3.1 Nowcast

$$\Omega_v = \left\{ x_{i,t_i}, t_i = 1, 2, \dots, T_{i,v}, i = 1, \dots, n; y_{3k}^Q, 3k = 3, 6, \dots, T_{Q,v} \right\} \quad (2)$$

e $T_{i,v}$ é o último período no qual os dados do período v a série i foi observada. Note que devido a não sincronia entre as séries, $T_{i,v}$ não é necessariamente o mesmo para todas as séries.

Uma das características do *nowcasting* é atualizar informação conforme os *releases* das variáveis são publicados. A primeira projeção será realizada com pouca ou quase nenhuma informação do trimestre. Em outras palavras, teremos uma sequência de projeções

$$E \left[y_t^Q | \Omega_v \right], E \left[y_t^Q | \Omega_{v+1}, \dots \right] \quad (3)$$

e $v, v + 1, \dots$ são datas de releases consecutivos. Tipicamente, os intervalos de tempo v são pequenos e frequentes. Agora, vamos esclarecer como e porque temos informação *nova* (vou chamar de *novidade* daqui para frente) e atualização das projeções e revisões do *nowcast* produzido. A utilização destas *novidades* de informação dentro do período considerado é de extrema relevância, e conforme citado em (Giannone, Reichlin e Small, (2006)) não temos motivos para desprezar quaisquer novas informações publicadas, não importa a frequência, pelos órgãos competentes (como, por exemplo, produção industrial e pesquisa de empregos para previsão do PIB).

Em resumo, podemos resumir nossos passos daqui para frente da seguinte maneira:

1. Atualizar a estimativa do trimestre atual e prever as variáveis de interesse, condicionadas ao largo conjunto de informações
2. Atualizar com base em um painel que é desbalanceado ao final do mês
3. Avaliar as *novidades* em relação à publicação das divulgações
4. Avaliar incertezas em relação ao fluxo de informação
5. Avaliar o impacto de uma nova publicação separando o efeito do tempo e da qualidade;

Veja que no período $v+1$ temos certo release de variáveis $\{x_{j,Tj,v+1}, j \in J_{v+1}\}$. O *novo* conjunto de informações difere do anterior pelos seguintes motivos. Primeiro, ele contém novos dados,

3.2 Modelagem Matemática

e segundo, porque poderemos ter revisões de dados antigos. Fato alíás bastante comum nas publicações de agências estatísticas brasileiras como o IBGE (Instituto Brasileiro de Geografia e Estatística) e outras pelo mundo.

Reescrevendo 3 em termos do termo *novo*, e assumindo que os dados são Normais, podemos achar os coeficientes $b_{j,t,v+1}$ tais que:

$$E \left[y_t^Q | \Omega_{v+1} \right] = E \left[y_t^Q | \Omega_v \right] + \sum_{j \in J_{v+1}} b_{j,t,v+1} \left(x_{j,T_{j,v+1}} - E \left[x_{j,T_{j,v+1}} | \Omega_v \right] \right) \quad (4)$$

sendo $E \left[y_t^Q | \Omega_{v+1} \right]$ é o novo *forecast*, $E \left[y_t^Q | \Omega_v \right]$ é o antigo *forecast* e $\left(x_{j,T_{j,v+1}} - E \left[x_{j,T_{j,v+1}} | \Omega_v \right] \right)$ é a *novidade*. Em outras palavras, podemos escrever a revisão de previsão como uma soma dos pesos dos novos dados publicados:

$$E \left[y_t^Q | \Omega_{v+1} \right] - E \left[y_t^Q | \Omega_v \right] = \sum_{j \in J_{v+1}} b_{j,t,v+1} \left(x_{j,T_{j,v+1}} - E \left[x_{j,T_{j,v+1}} | \Omega_v \right] \right) \quad (5)$$

Sendo que o tamanho da revisão depende do tamanho da *novidade*.

3.2 Modelagem Matemática

Para a estimação dos modelos de *nowcasting*, poderíamos utilizar um arcabouço mais complexo de estimação, como modelos fatoriais ou *mixed data sampling (MIDAS)*, por exemplo, que são maneiras parsimoniosas mas restritivas de combinar um conjunto extenso de informação com dados de frequência mistos e com a presença de *ragged edges*. Entretanto, os resultados são de difícil interpretação. Desta maneira, preferimos uma abordagem mais simplificada e extensamente utilizada na literatura para estimação de modelos de *nowcasting*: vamos utilizar o conceito de *bridge equations*, que relaciona o PIB à indicadores mensais agregados de uma ou mais séries. Segundo (Castle, Hendry e Kitov, (2013)), esta abordagem é inclusive utilizada pelos Bancos Centrais de muitos países devido à sua facilidade de interpretação. Esta abordagem estima uma equação para a variação do PIB, Δy_{t_q} com uma frequência trimestral:

3.2 Modelagem Matemática

$$\Delta y_{t_q} = c + \sum_{i=1}^p \alpha_i \Delta y_{t_{q-1}} + \sum_{j=0}^p \sum_{i=1}^k \beta_{ij} z_{i,t_q-j} + u_{t_q} \quad (6)$$

para $t = 1, \dots, T$

sendo que k indicadores trimestrais z_{i,t_q} são transformados em estacionário, p é o lag e assume-se que u_t são resíduos IID. Veja que os parâmetros desta regressão descrevem uma relação entre os indicadores e o PIB trimestral.

Para comparação de desempenho deste conjunto de regressões, vamos avaliar os resultados em relação a um modelo benchmarking AR (4) para a série trimestral com sazonalidade do PIB do Brasil obtido no website do IBGE. Tal modelo foi utilizado após análise do correlograma da série mostrar ser este o melhor modelo conceitual ajustado aos dados. Desta maneira, estamos estimando a equação:

$$y_t^Q = c + \sum_{i=1}^4 \beta_i y_{t-i}^Q + \varepsilon_t^Q \quad (7)$$

sendo c a constante do modelo, ε_t^Q é o termo ruído branco trimestral tal que $\varepsilon_t^Q \sim N(0, \sigma_\varepsilon^2)$.

4 Seleção automática de modelos

Conforme citado nas sessões anteriores, dado o arcabouço de modelos que estamos trabalhando, precisamos de um sistema de seleção automática de modelos que nos auxilie quando temos mais variáveis que observações, permitindo trabalhar com todas as informações disponíveis, e tratamento de múltiplas quebras estruturais. Conforme a abordagem de (Castle, Hendry e Kitov, (2013)), podemos utilizar o *Autometrics* com *impulse saturation* conforme apresentado em (Castle, Doornik e Hendry, (2012)). A teoria parte da existência de um processo de geração de dados local (da sigla LDGP em inglês) para qualquer subconjunto de variáveis, e a relação de qualquer modelo àquele LDGP. A ideia por trás da seleção de modelos geral para específico é localizar uma boa aproximação ao LDGP. Para isto, o algoritmo inicia com o modelo mais geral possível contendo todas as variáveis candidatas, suas defasagens transformações na forma funcional e possíveis quebras, o que nos leva em geral ao problema de mais variáveis, N , que observações T .

A abordagem então é construir um modelo com o conjunto de dados baseados em pressupostos teóricos gerais que contém o LDGP, e a partir deste conjunto, reduzir o modelo de sua forma geral para uma representação específica e parcimoniosa. Este é um procedimento em duas etapas:

1. Definir um conjunto de N variáveis que inclui a LDGP como um submodelo
2. Iniciando com o modelo geral como uma boa aproximação para todas as propriedades do dado, reduzir sua complexidade removendo variáveis insignificantes, enquanto avalia-se a cada redução a validade do modelo geral é preservada.

Este é o conceito básico do modelo geral para específico (da sigla *gets* em inglês).

4.1 *Autometrics*

O conceito apresentado anteriormente da metodologia *Gets* está disponível no software *Autometrics*, disponível no pacote *Oxmetrics 6.0*. O algoritmo está baseado em cinco componentes:

4.2 *Impulse-Indicator Saturation (IIS)*

1. GUM: o modelo geral irrestrito (sigla GUM em inglês) é o ponto inicial da busca, e deve ser formulado baseado em considerações teóricas embasadas, pois deve conter o LGDP.
2. Pré-Busca: antes da seleção em si, uma pré-busca para redução de defasagens é desenvolvida, para remover defasagens insignificantes, aumentando o desempenho na busca por variáveis relevantes. Somente é aplicada se o número de variáveis não exceder o número de observações ($N > T$).
3. Caminhos de busca: o *Autometrics* utiliza uma árvore para explorar caminhos. Iniciando com o GUM, o algoritmo elimina a variável menos significativa determinado pela menor razão t. Cada remoção é um ramo desta árvore. Para cada redução, há uma única sub-árvore e cada remoção é backtestada contra o modelo GUM inicial através de um teste F. Se este falha, nenhum sub-nó deste ramo é considerado. Este passo é executado até que nenhuma variável possa ser removida no nível de significância α previamente escolhido. Neste caso, o modelo é considerado terminal.
4. Teste Diagnóstico: cada modelo terminal é submetido a uma série de testes dado um nível de significância: Normalidade, Heterocedasticidade, teste de Chow e correlação residual e heterocedasticidade autoregressiva condicional.
5. Desempate: como resultado da busca em árvores, múltiplos modelos terminais podem ser encontrados. Para desempate, o critério de Schwarz é utilizado.

A calibragem principal do algoritmo de busca é a escolha do nível de significância α . Como resultado deste trabalho, veremos a importância da escolha deste parâmetro, e os diferentes resultados que podemos obter com diferentes escolhas.

4.2 **Impulse-Indicator Saturation (IIS)**

Segundo (Castle, Doornik e Hendry, (2012)) a abordagem para tratar e remover múltiplas quebras estruturais das séries é incluir um indicador de impulso para cada observação do conjunto

4.3 Teste de Diebold-Mariano (DM) para comparar previsões

de regressores candidatos, chamado de saturação de indicadores de impulso (da sigla IIS - *Impulse-indicator saturation* em inglês). Quanto T variáveis de impulso de indicadores são combinadas com os outros regressores candidatos, o seu número total deve exceder o tamanho da amostra. Johansen e Nielsen provaram que sob a hipótese nula de nenhum *outlier* ou *shift* não existe quase nenhuma perda de eficiência em testar T indicadores de impulso para $\alpha \leq \frac{1}{T}$, mesmo em modelos dinâmicos. Aparentemente curioso à primeira vista, reter um indicador de impulso quando não é necessário apenas remove uma observação, o que é tudo que acontece em média. Logo, a eficiência é da ordem de $100(1 - \alpha)\%$.

Sendo assim, o IIS utiliza indicador de *dummies* 0-1 de impulso $\{I_{it}\}$ para analisar as propriedades de um modelo. Para uma amostra de T observações, há uma quantidade T destas *dummies*, então a inclusão irrestrita de todas as T destas *dummies* no modelo (saturando a amostra) pode torná-lo infactível. Entretanto, blocos de *dummies* podem ser incluídos, e esta característica é a base da proposta do IIS.

Conforme apresentado na seção 2 é de extrema relevância a utilização desta técnica para a execução de *nowcasting* robusto.

4.3 Teste de Diebold-Mariano (DM) para comparar previsões

Para comparação da eficácia de previsão dos modelos que serão utilizados neste trabalho, vamos utilizar o teste de *Diebold-Mariano*.

Considerando duas séries objetivas de previsão $y_{t+h|t}^1$ e $y_{t+h|t}^2$, suponha que tenhamos dois *forecasts* de y_{t+h} . O erro de *forecast* será então:

$$\varepsilon_{t+h|t}^1 = y_{t+h} - y_{t+h|t}^1 \quad (8)$$

$$\varepsilon_{t+h|t}^2 = y_{t+h} - y_{t+h|t}^2 \quad (9)$$

o h -ésimo *forecast* será computado para $t = t_0, \dots, T$ para um total de T_0 *forecasts* resultando:

$$\left\{ \varepsilon_{t+h|t}^1 \right\}_{T_0}^T, \left\{ \varepsilon_{t+h|t}^2 \right\}_{T_0}^T \quad (10)$$

4.3 Teste de Diebold-Mariano (DM) para comparar previsões

Como o *forecast* do h -ésimo passo utiliza dados sobrepostos os erros de *forecast* em $\left\{ \varepsilon_{t+h|t}^1 \right\}_{T_0}^T$ e $\left\{ \varepsilon_{t+h|t}^2 \right\}$ serão correlacionados.

A eficácia de cada *forecast* é medida por uma função perda particular:

$$L(y_{t+h}, y_{t+h|t}^i) = L(\varepsilon_{t+h|t}^i), \quad i = 1, 2 \quad (11)$$

Algumas funções perda que podem ser utilizadas são:

- Erros quadrados: $L(\varepsilon_{t+h|t}^i) = \left(\varepsilon_{t+h|t}^i \right)^2$
- Erros absolutos: $L(\varepsilon_{t+h|t}^i) = \left| \varepsilon_{t+h|t}^i \right|$

Para determinar se um modelo realiza previsões melhores que outro, podemos testar a seguinte hipótese nula:

$$H_0 : E[L(\varepsilon_{t+h|t}^1)] = E[L(\varepsilon_{t+h|t}^2)] \quad (12)$$

contra a seguinte hipótese alternativa:

$$H_1 : E[L(\varepsilon_{t+h|t}^1)] \neq E[L(\varepsilon_{t+h|t}^2)] \quad (13)$$

O teste de Diebold-Mariano é baseado na função perda:

$$d_t = L(\varepsilon_{t+h|t}^1) - L(\varepsilon_{t+h|t}^2) \quad (14)$$

Sendo assim, a hipótese nula de acurácia preditiva é então:

$$H_0 : E[d_t] = 0 \quad (15)$$

A estatística de teste Diebold-Mariano é a seguinte:

$$S = \frac{\bar{d}}{(\widehat{avar}(\bar{d}))^{1/2}} = \frac{\bar{d}}{\left(\widehat{LRV}_{\bar{d}}/T \right)^{1/2}} \quad (16)$$

4.3 Teste de Diebold-Mariano (DM) para comparar previsões

sendo:

$$\bar{d} = \frac{1}{T_0} \sum_{t=t_0}^T d_t \quad (17)$$

$$LRV_{\bar{d}} = Y_0 + 2 \sum_{j=1}^{\infty} Y_j, Y_j = cov(d_t, d_{t-j}) \quad (18)$$

e $\hat{LRV}_{\bar{d}}$ é uma estimativa consistente da variância assintótica (longo prazo) de $\sqrt{T}\bar{d}$. A variância longo prazo é utilizada na estatística porque os erros $\{d_t\}_{T_0}^T$ são correlacionados para $h > 1$. Diebold-Mariano mostraram que sob a hipótese nula de acurácia preditiva:

$$S^a \sim N(0, 1) \quad (19)$$

Logo rejeitamos a hipótese nula de acurácia igual no nível de 5% se

$$|S| > 1,96$$

5 Metodologia

5.1 Base de Dados

Sabe-se historicamente que a economia brasileira possui instabilidades inerentes à suas características, com múltiplas quebras estruturais em sua série do PIB. (Valls e Vieira, (2014)) fizeram um estudo detalhado dos ciclos de negócio brasileiro desde o ano de 1900 até 2012. Na década de 1980, conhecida como a “década perdida” apresenta, por exemplo, 24 trimestres de recessão e 16 de expansão. Neste período também tivemos aumento nos preços internacionais do petróleo, e taxas de juros, e um grande peso de países endividados na América Latina, culminando com a moratória do México em 1982, o que levou à restrição de crédito dos bancos internacionais na região, agravando a situação de liquidez dos países desta região. Neste período, o Brasil passou por uma fase de crescimentos pequenos, entre 1981 e 1983. No final da década de 1980 e início de 1990 a economia brasileira passou por muitas turbulências, processo hiperinflacionário e 6 planos de estabilização para tentar contê-lo, o que causou quebras estruturais na economia e acabou criando incertezas e instabilidade no cenário econômico brasileiro.

Por outro lado, a década de 1990 apresentou 20 trimestres de recessão e 20 trimestres de expansão, com destaque para o plano Collor (março de 1990) trouxe 10 trimestres de recessão e uma sequência de crises financeiras entre 1995 e 1999 trouxeram 13 trimestres de recessão. Além disto, nesta década o Brasil teve dois grandes períodos de depressão: um após o plano Cruzado de 1987 a 1993 e outro após o plano Real de 1995 a 1999. Em resumo, tivemos ao longo dos anos uma série de planos e mudanças na política econômica ao longo dos anos para tentar a estabilização da economia, o que aumenta a necessidade de tratamento de quebras estruturais quando consideramos modelos de previsão do PIB brasileiro. A figura 1 mostra a evolução do PIB trimestral medido pelo IBGE ao longo dos anos a partir de 1980, com a representação de sua série em nível e primeira diferença em logaritmo, sua distribuição e o respectivo correlograma indicando a memória desta série.

5.1 Base de Dados



Figura 1: Representações gráficas da série do PIB trimestral do Brasil - fonte: Ipeadata

Vamos utilizar o algoritmo *Autometrics* para seleção de variáveis, e buscamos algumas referências de trabalhos acadêmicos para avaliar quais variáveis e indicadores foram importantes em algum momento da história para explicar a variação do PIB Brasileiro, com isto, esperamos cobrir grande parte deste universo. Neste sentido, os trabalhos de (*Chauvet, (2000)*), (*Ferreira, (2005)*) e (*Chauvet, Elcyon e Vasquez, (2002)*) foram de extrema importância para coleta de bons indicadores e idéias de boas variáveis predictoras da variação do PIB no caso brasileiro.

Ao todo, foram selecionadas 72 variáveis da economia nacional e internacional, subdivididas nas categorias apresentadas na tabela 1.

5.1 Base de Dados

Tabela 1: Indicadores mensais utilizados como variáveis explicativas nos modelos

Tipo	Grupo	Qt. de indicadores
Real	Produção Industrial	17
Real	Produção de automóveis	6
Real	Mercado de Trabalho	4
Real	Consumo de Energia Elétrica	5
Real	Produção Industrial - Demais Países	5
Financeira	Mercado Financeiro	10
Financeira	Mercado Financeiro - EUA	5
Financeira	Balança Comercial - Brasil	12
Financeira	Balança Comercial - Demais Países	4
Commodities	Preços e Produção	4
TOTAL		72

Fonte: elaboração do próprio autor.

Alguns comentários a respeito do grupo inicial das variáveis utilizadas:

- Todas as variáveis são de periodicidade mensal, exceto o PIB do Brasil que é trimestral;
- Sabemos da possível relação da economia brasileira com a economia de países industrializados e commodities, por isto utilizamos indicadores nestes grupos também;
- Neste trabalho utilizamos o período desde jan-80 a dez-14. Por este motivo, alguns indicadores não puderam ser utilizados, como, por exemplo, indicadores de confiança e expectativas. Sabemos da importância destes indicadores na explicação da variação do PIB brasileiro, e que poderíamos utilizar um modelo para completar os dados faltantes. Entretanto, preferimos não incluir tais variáveis no modelo, dado o universo de variáveis consideradas.
- As séries de produção industrial do primeiro bloco sofreram alterações de metodologia ao longo do tempo pelo IBGE, portanto, tiveram que ser encadeadas no período considerado;

5.1 Base de Dados

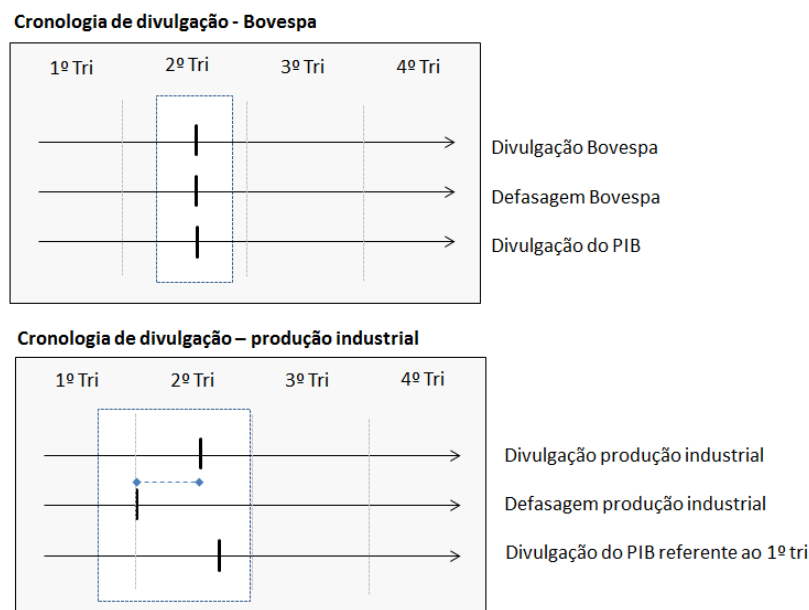


Figura 2: Exemplo de defasagens de publicação dos indicadores - IBOVESPA e produção industrial.

- Defasagens: Note que conforme já citado por outros trabalhos apresentados nas seções anteriores, os diversos indicadores são publicados em períodos diferentes dentro do mês, com diferentes defasagens. Quando estamos falando de *nowcasting*, precisamos considerar este fato. Na figura 2 vemos um exemplo deste item, com a publicação dos indicadores do IBOVESPA e produção industrial. Veja que para o primeiro, temos a informação mais atualizada no momento de divulgação do PIB, fato que não é observado na segunda variável - temos divergências entre a divulgação e a defasagem. No apêndice e na figura 9 são exibidas todas as variáveis agrupadas por período de publicação e defasagens, para um trimestre qualquer. Note que se considerarmos o primeiro mês do trimestre, teremos indicadores saindo na primeira e segunda quinzena, porém, contendo informações diferentes de defasagem conforme pode ser visto pelas diferentes cores exibidas. Esta característica deverá ser modelada e tratada conforme a modelagem apresentada nas seções anteriores.
- Algumas das variáveis são $I(0)$, $I(1)$ ou $I(2)$, porém seguindo a recomendação de (Castle, Hendry e Kitov (2013)) fizemos as transformações necessárias e os testes de raiz unitária para garantir que todas fossem estacionárias, e trabalhamos com a diferença dos logs.

5.2 Especificação dos modelos

O gráfico de algumas destas variáveis encontra-se na figura 3, e temos respectivamente: produção industrial - bens de consumo, indústrias extrativas, metalurgia básica, outros produtos químicos e indústria têxtil, horas pagas na indústria de SP, exportações (FOB) e exportações quantum. Note que na série de produção industrial - extrativa mineral quantum (DLPI_EXTR_MIN) temos em 1995 um claro *outlier* positivo. Após a implantação do plano real, a produção extrativa mineral brasileira apresentou um forte crescimento durante 1994, com acentuado crescimento na demanda, fato que se estendeu-se durante o ano de 1995. Após este período, observamos uma forte queda neste indicador.

- No apêndice mostramos a lista completa de variáveis que foram utilizadas na figura 8.

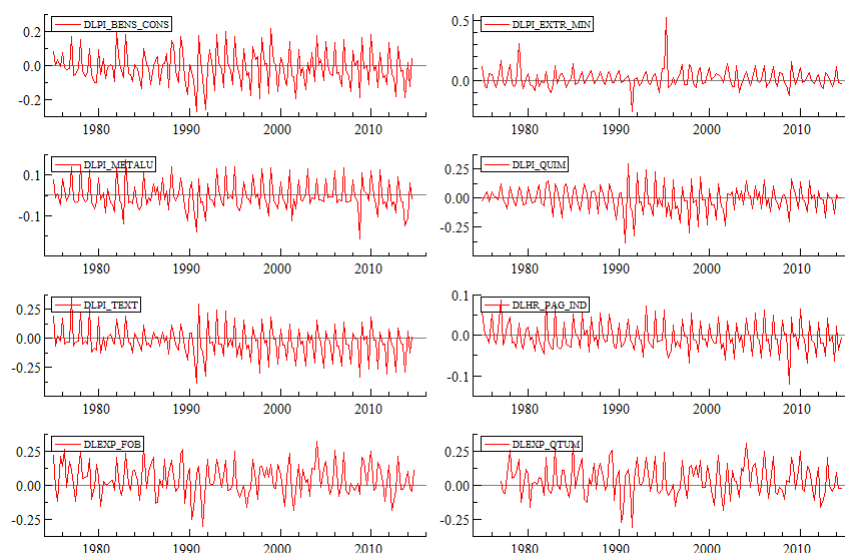


Figura 3: Gráfico mostrando a estacionariedade de algumas das variáveis utilizadas nos modelos deste trabalho

5.2 Especificação dos modelos

Com a base de dados estruturada, passamos então ao desenvolvimento e rodada dos modelos realizada no software *Oxmetrics* 6.0. A montagem e estruturação dos modelos foi feita levando-se em conta os pontos apresentados no referencial teórico, de maneira a capturarmos as visões

5.2 Especificação dos modelos

necessárias para comparação dos modelos. Desta maneira, a estruturação foi feita segundo a tabela 2 abaixo.

Método	Modelo	Defasagem das variáveis	Defasagem do PIB
Nowcast	NC3	0 a 6	1 a 6
	NC2	1 a 7	
	NC1	2 a 8	
Forecast	F3	3 a 9	2 a 7
	F2	4 a 10	3 a 8
	F1	5 a 11	4 a 9

Tabela 2: Modelos que serão utilizados neste trabalho, com a especificação de suas respectivas variáveis explicativas.

Desta maneira, temos as seguintes especificações de modelos neste trabalho:

1. Modelo 1: Simulação de um modelo de *nowcasting* (NC3) - modelo considerando a defasagem de 0 a 6 de todos os indicadores mensais e as defasagens de 1 a 6 do log(PIB). Este modelo contempla a publicação completa dos indicadores mensais (release 3) dentro do trimestre.
2. Modelo 2: Simulação de um modelo de *nowcasting* (NC2) - modelo considerando a defasagem de 1 a 7 de todos os indicadores mensais e as defasagens de 1 a 6 do log(PIB). Este modelo contempla a publicação de 2 meses dos indicadores mensais (release 2) dentro do trimestre.
3. Modelo 3: Simulação de um modelo de *nowcasting* (NC1) - modelo considerando a defasagem de 2 a 8 de todos os indicadores mensais e as defasagens de 1 a 6 do log(PIB). Este modelo contempla a publicação de 1 mês dos indicadores mensais (release 1) dentro do trimestre.
4. Modelo 4: Simulação de um modelo de *forecasting* (F3) - modelo considerando a defasagem de 3 a 9 de todos os indicadores mensais e as defasagens de 2 a 7 do log(PIB).

5.2 Especificação dos modelos

5. Modelo 5: Simulação de um modelo de *forecasting* (F2) - modelo considerando a defasagem de 4 a 10 de todos os indicadores mensais e as defasagens de 3 a 8 do log(PIB).
6. Modelo 6: Simulação de um modelo de *forecasting* (F1) - modelo considerando a defasagem de 5 a 11 de todos os indicadores mensais e as defasagens de 4 a 9 do log(PIB).

A proposta da utilização e especificação destes modelos é proporcionar medição e comparação dos erros de previsão ao longo do trimestre, e atualização conforme novos indicadores são publicados pelos órgãos competentes. Conforme referencial teórico, diversos trabalhos empíricos comprovaram os ganhos de desempenho de previsão dos modelos com a incorporação das novidades de *releases* mensais, fato que gostaríamos de testar com os dados no caso do Brasil.

Para cada um dos modelos, avaliamos os critérios de *backtest* para períodos de 10 (2004-2014) e 15 anos (1999-2014), porém, acabamos por escolher este último período para o trabalho devido à maior quantidade de pontos para análise e comparação dos resultados. Para comprovação dos resultados, efetuamos em seguida a comparação entre os modelos através do teste de Diebold-Mariano, conforme descrito na seção 4.3.

Em resumo, a proposta foi testar estatisticamente a qualidade de previsão dos modelos NC3, NC2, NC1, F3, F2, F1 utilizando o arcabouço de *bridge equations* e tendo como *benchmark* um modelo AR (4) sem variáveis explicativas. Também para efeitos de comparação, testamos dois diferentes níveis de significância de 1% e 0,1% na rodada do modelo e seleção de variáveis no *Autometrics*. Conforme os resultados que serão apresentados logo abaixo, este último fator mostrou-se preponderante para a diferenciação dos resultados obtidos. Além disto, mostramos empiricamente que este mesmo resultado se aplica quando analisamos o desempenho dos modelos quando levamos em conta os ciclos econômicos da economia brasileira.

6 Resultados

6.1 Modelos

Após o desenvolvimento e rodada dos modelos NC3, NC2, NC1, F3, F2, F1 no software *Oxmetrics* 6.0, obtivemos o resultado geral apresentado na figura 4. A proposta foi dividir a análise através de duas variáveis importantes de entrada do modelo: a significância e a presença de variáveis *dummies* sazonais. Lembramos que conforme (Castle, Hendry e Kitov, (2013)), esperamos que a utilização de variáveis *dummies* aumente o desempenho de previsão dos modelos.

O gráfico da figura 4 resume os resultados obtidos com a rodada de cada um dos modelos de *forecast* e *nowcast*, com e sem variáveis *dummies* (*D*), e variando para significâncias 1% e 0,1%.

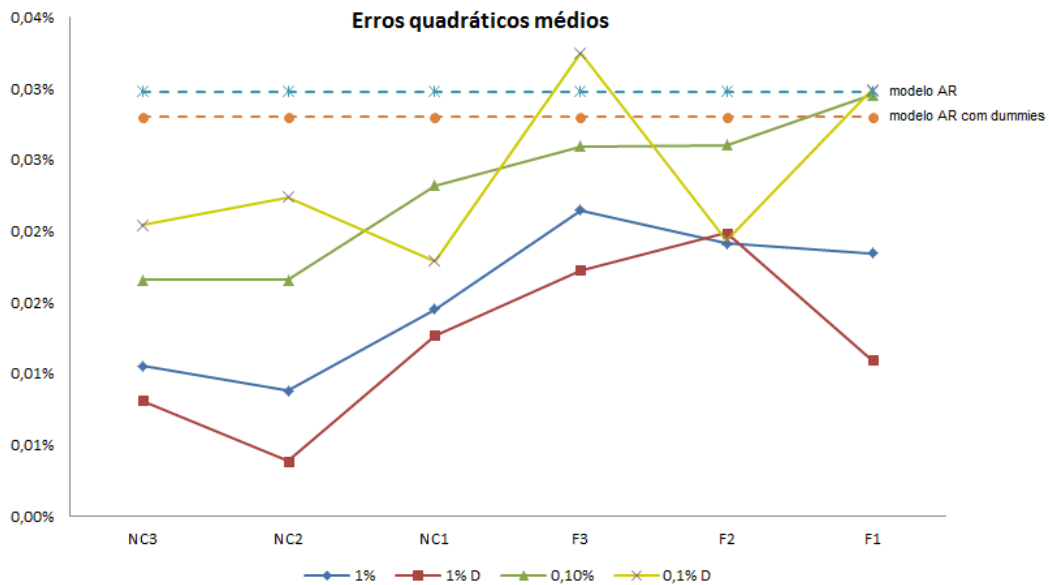


Figura 4: Comparação do EQM para os modelos.

Em seguida vamos discutir algumas conclusões que podemos tirar a respeito destes resultados².

Com uma visualização geral do gráfico, veja que o modelo AR utilizado de *benchmark* apresentou erros de previsão acima dos outros modelos, exceto para o modelo F3 quando utilizamos variáveis *dummies* com significância ($\alpha = 0,1\%$). Este é um resultado importante e esperado,

²Importante: todas as afirmações aqui citadas foram testadas estatisticamente ao longo do trabalho e serão apresentadas mais a frente.

6.1 Modelos

que está conforme o trabalho de (Castle, Hendry e Kitov, (2013)).

Quando comparamos as performances através da significância utilizada na seleção, para todos os modelos, exceto para o modelo F2 que temos um empate, a utilização da significância ($\alpha = 1\%$) é estritamente melhor que a utilização da significância ($\alpha = 0,1\%$) considerando os modelos com ou sem utilização de variáveis *dummies*. No gráfico, as linhas em azul e vermelho estão com erros sempre menores que as demais. Este não era um resultado esperado, e nos parece ser uma característica especial dos dados da economia brasileira (dois primeiros gráficos da linha superior na figura 5).

Quando comparamos os desempenhos através da utilização de variáveis *dummies*, notamos que para a significância ($\alpha = 1\%$), vemos que a utilização de variáveis *dummies* realmente melhora a qualidade das previsões dos modelos (no gráfico, comparação entre as linhas azul e vermelha). Já no caso da significância ($\alpha = 0,1\%$) vemos que este resultado não se repete (no gráfico, comparação entre as linhas roxa e amarela), mostrando novamente que a significância é importante neste contexto (dois gráficos da linha inferior na figura 5).

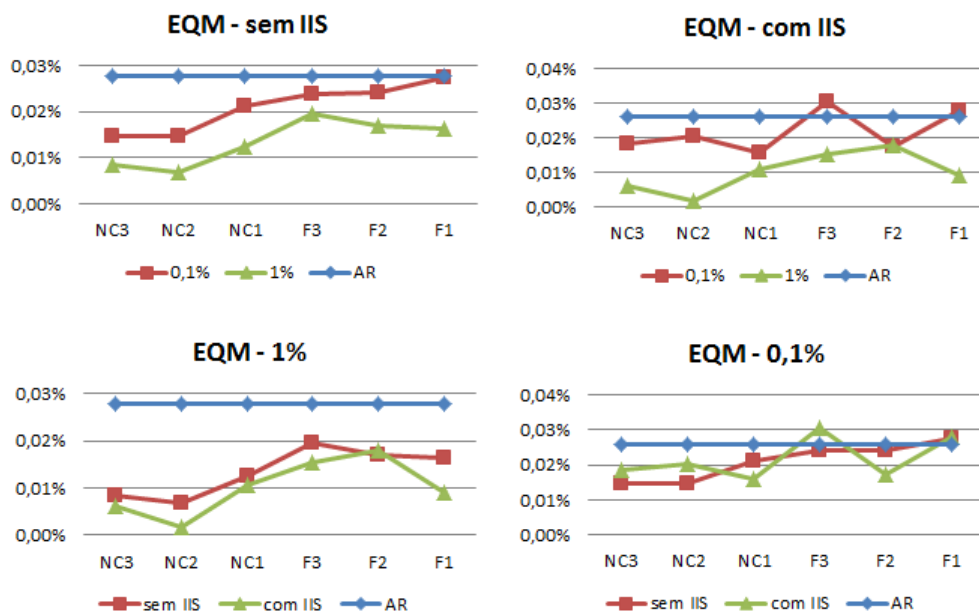


Figura 5: Gráfico com os erros de previsão por modelo.

Por último, vemos uma tendência crescente dos erros conforme vamos evoluindo nos modelos NC3, NC2, NC1, F3, F2, F1, mostrando que a utilização do conceito de *nowcast* aumenta o desempenho de previsão em relação ao conceito de *forecast* puro, conforme as publicações dos indicadores vão acontecendo ao longo dos meses dentro do trimestre. Veja que em geral, os

6.1 Modelos

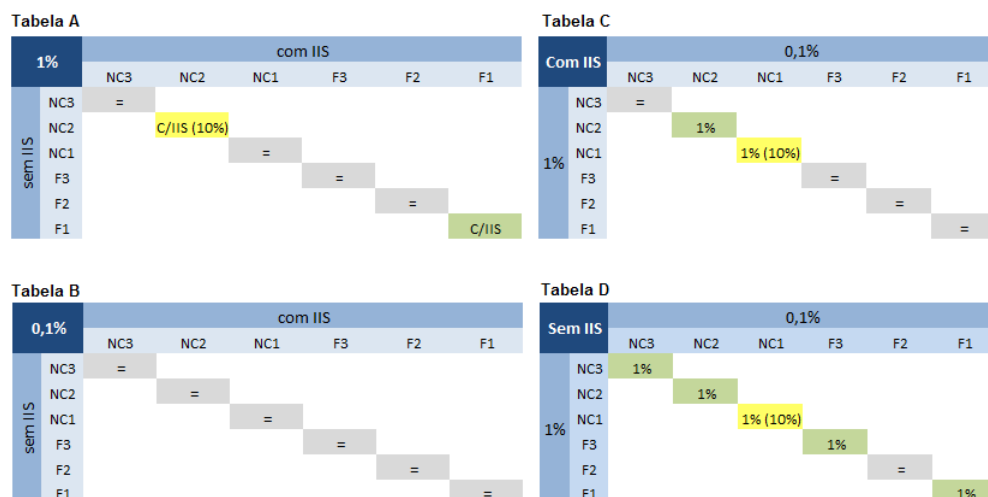


Figura 6: Testes de Diebold-Mariano

erros para o modelo completo NC3 (simulando a publicação completa das variáveis) é menor que os demais modelos na maioria dos casos.

Em geral, os testes confirmaram estatisticamente as suposições feitas ao analisarmos os gráficos de resultados com os erros de previsão. A tabela da figura 6 mostra os resultados obtidos com este teste.

Para cada modelo, a tabela mostra o resultado da regressão obtida através do teste de Diebold-Mariano. As tabelas à esquerda mostram o resultado dos testes para a significância da seleção de variáveis em 1% e 0,1% respectivamente. As tabelas à direita mostram os resultados fixando-se a utilização de variáveis *dummies*. As cores denotam: igualdade nos resultados (cinza), ganhos em relação ao desempenho com 5% de significância do teste DM (verde) e 10% no teste DM (amarelo).

É possível notar que:

1. Fixando-se a significância do modelo de seleção de variáveis **em 1%** (tabela A), a utilização de variáveis *dummies* mostrou-se melhor para os modelos NC2 (com 10% de significância no teste) e o modelo F1. Para todos os demais, a utilização de variáveis *dummies* não fez diferença para melhorar a eficácia das previsões.
2. Fixando-se a significância do modelo de seleção de variáveis **em 0,1%** (tabela B), a utilização de variáveis *dummies* não fez diferença para melhorar a eficácia das previsões, em nenhum dos modelos analisados.

6.1 Modelos

3. Fixando-se os modelos **com uso** de variáveis *dummies* (tabela C), vemos que a significância utilizada no modelo de seleção mostrou-se melhor nos modelos NC2 e NC1 (com 10% de significância no teste) quando utilizamos este parâmetro em 1%. Para todos os demais, a significância não fez diferença para melhorar a eficácia das previsões.
4. Fixando-se os modelos **sem uso** de variáveis *dummies* (tabela D), vemos que a significância utilizada no modelo de seleção mostrou-se melhor na maioria dos casos para os modelos NC3, NC2, NC1 (com 10% de significância no teste), F3 e F1 quando utilizamos este parâmetro em 1%. Unicamente para o modelo F2 este parâmetro não fez diferença.

Como resultado geral, vemos que a significância de 1% mostrou-se melhor que o valor de 0,1% na maioria dos casos, porém, não podemos concluir que a utilização de variáveis *dummies* fez diferença no horizonte de avaliação deste trabalho. Este resultado para a economia brasileira é um pouco diferente do trabalho de (Castle, Hendry e Kitov, (2013)), porém, entendemos ser um resultado significativo que descreve as características da dinâmica desta economia. Intuitivamente, não era um resultado esperado termos um melhor desempenho quando utilizamos um critério de seleção menos apertado na seleção de variáveis, no entanto, entendemos que no caso dos dados brasileiros, isto possibilitou reter variáveis que se mostraram mais importantes na previsibilidade do PIB, e que não foi possível quando apertamos este indicador.

Os resultados do teste de Diebold-Mariano para comparação com o modelo *benchmarking* estão na tabela da figura 7. As colunas representam o resultado do teste para os modelos, de acordo com a significância e a utilização ou não de variáveis *dummies* (IIS). Quando comparamos os modelos NC3, NC2, NC1, F3, F2, F1 ao modelo AR utilizado como *benchmark*, vemos que novamente diferenças de resultado conforme mudamos a significância. Por exemplo, a coluna 1% com IIS refere-se aos modelos com 1% de significância e utilização de variáveis *dummies*. Para a significância de 1%, vemos que os modelos NC3, NC2, NC1, F3, F2, F1 venceram o modelo AR de benchmark em todas as situações, exceto para o caso do F3. Já no caso da significância de 0,1% vemos resultados dúbios, e temos igualdade de resultados entre o *benchmark* e os modelos NC2, F3, F2 e F1 para o caso sem *dummies* (IIS), e NC2, F3 e F1 para o caso com *dummies* (IIS).

6.2 Ciclos Econômicos

Modelo	1%	1% com IIS	0,1%	0,1% com IIS
NC3	NC3	NC3	NC3	NC3
NC2	NC2	NC2	=	=
NC1	NC1	NC1	NC1	NC1 (10%)
F3	=	F3	=	=
F2	F2	F2	=	F2
F1	F1	F1	=	=

Figura 7: Teste de Diebold-Mariano para comparação com o modelo utilizado como *benchmarking*.

6.2 Ciclos Econômicos

Outra análise realizada que reforça a importância da avaliação de quais significâncias adotar foi a relação de desempenho dos modelos de previsão em relação ao ciclo de expansão na qual a economia brasileira se encontra. (Valls e Vieira, (2014)) mostraram a datação dos ciclos de negócio brasileiro de 1900 até 2012. Para efeitos de comparação e eficácia do modelo de *nowcasting*, vamos analisamos o desempenho dos modelos de acordo com a classificação deste trabalho no período de avaliação dos modelos. A tabela 3 relaciona a média da somatória ao quadrado dos erros dos 6 modelos (NC1, NC2, NC3, F3, F2, F1). As duas primeiras colunas utilizam significância ($\alpha = 1\%$) e as duas últimas utilizam ($\alpha = 0,1\%$). A segunda e a quarta colunas referem-se aos modelos com utilização de variáveis *dummies*.

Período	Ciclo	Média dos erros de previsão ao quadrado			
		(1%)	(1%) c/ IIS	(0,1%)	(0,1%) c/ IIS
2000(1) a 2001(1)	Expansão	1,26E-05	2,36E-07	6,93E-05	7,17E-07
2001(2) a 2003(4)	Recessão	5,28E-07	2,57E-07	1,01E-06	3,76E-05
2004(1) a 2008(3)	Expansão	7,52E-05	6,73E-08	2,60E-07	2,50E-07
2008(4) a 2009(3)	Recessão	1,20E-06	1,03E-06	1,62E-06	3,05E-06
2009(4) a 2011(2)	Expansão	3,41E-05	2,63E-08	4,85E-07	2,46E-05
2011(3) a 2012(4)	Recessão	5,80E-08	5,57E-08	3,58E-06	9,07E-06

Tabela 3: Resultado do EQM dos 6 modelos por ciclo econômico.

Fonte: elaboração do próprio autor.

Como exemplo, a tabela 4 apresenta os resultados para o modelo NC3 temos (demais modelos apresentados no apêndice). A tabela da esquerda apresenta os resultados para significância

6.3 Variáveis finais do modelo

de ($\alpha = 1\%$) para a seleção de variáveis, enquanto que a tabela à direita apresenta os resultados para ($\alpha = 0,1\%$)

Ciclo econômico	sem IIS	com IIS	%	Ciclo econômico	sem IIS	com IIS	%
Expansão	1,27E-04	9,71E-05	76%	Expansão	2,45E-04	1,95E-04	80%
Recessão	1,17E-04	5,75E-05	49%	Recessão	1,04E-04	2,03E-04	196%
Relação	92%	59%		Relação	42%	104%	

Tabela 4: EQM do modelo NC3 por ciclo econômico.

Fonte: elaboração do próprio autor.

Neste caso, para a significância de 1%, observando a coluna % vemos a média dos erros utilizando IIS foi melhor tanto aos períodos de expansão quanto de recessão. Da mesma maneira, os erros foram menores para o período de recessão quando considerados modelos com ou sem uso de variáveis *dummies*.

No caso da significância de 0,1%, observando a coluna % vemos a média dos erros utilizando IIS foi melhor tanto no período de expansão, e pior no período de recessão. Da mesma maneira, os erros foram menores para o período de recessão quando considerados modelos sem uso de variáveis *dummies*, e pior no caso de modelos com variáveis *dummies*. Assim, estes resultados ambíguos para o Brasil quando variamos a significância dos modelos mostraram mais uma vez a importância da escolha deste parâmetro.

6.3 Variáveis finais do modelo

Exemplificando, vamos demonstrar os resultados obtidos da regressão utilizando o *autometrics* para o modelo NC3 com significância de 1% sem variáveis *dummies*.

6.3 Variáveis finais do modelo

$$\begin{aligned}
 \mathbf{DLPIB} = & 0.172 \mathbf{DLPIB}_{t-4} - 0.0497 \mathbf{DLPI_BENS_CAP}_{t-2} + 0.04188 \mathbf{DLPI_EXTR_MIN}_{t-2} \\
 & (0.0442) \qquad\qquad\qquad (0.0161) \qquad\qquad\qquad (0.0171) \\
 & + 0.073 \mathbf{DLPI_EXTR_MIN}_{t-5} - 0.2904 \mathbf{DLPI_IND_GER}_t + 0.2535 \mathbf{DLMAQ}_t \\
 & (0.0152) \qquad\qquad\qquad (0.0538) \qquad\qquad\qquad (0.0291) \\
 & - 0.1221 \mathbf{DLMAQ_EQ}_{t-6} - 0.04339 \mathbf{DLTRANSPO}_{t-6} + 0.1234 \mathbf{DLPI_ALIMEN}_t \\
 & (0.0246) \qquad\qquad\qquad (0.0133) \qquad\qquad\qquad (0.0224) \\
 & - 0.1189 \mathbf{DLPI_ALIMEN}_{t-1} + 0.1808 \mathbf{DLPI_PAPEL}_{t-6} + 0.5585 \mathbf{DLUTIL_CAP_INDSP}_{t-1} \\
 & (0.0187) \qquad\qquad\qquad (0.0335) \qquad\qquad\qquad (0.0666) \\
 & - 0.3378 \mathbf{DLHR_PAG_IND}_{t-4} + 0.3576 \mathbf{DLHR_PAG_IND}_{t-6} + 0.0126 \mathbf{DLSELIC}_{t-6} \\
 & (0.0546) \qquad\qquad\qquad (0.0691) \qquad\qquad\qquad (0.00404) \\
 & - 0.04633 \mathbf{DLM0}_{t-1} + 0.05218 \mathbf{DLM0}_{t-2} + 0.03506 \mathbf{DLM1}_{t-4} \\
 & (0.0092) \qquad\qquad\qquad (0.0087) \qquad\qquad\qquad (0.00832) \\
 & - 0.01485 \mathbf{DLEXEC_FIN_DESP}_{t-4} - 0.2458 \mathbf{DLEXP_FOB}_{t-2} + 0.2909 \mathbf{DLEXP_QTUM}_{t-2} \\
 & (0.00453) \qquad\qquad\qquad (0.0566) \qquad\qquad\qquad (0.057) \\
 & + 0.008731 \mathbf{DLAUTO_ONI_EXP}_{t-4} + 0.01608 \mathbf{DLMAQ_EXP}_{t-2} + 0.01807 \mathbf{DLMAQ_EXP}_{t-5} \\
 & (0.00284) \qquad\qquad\qquad (0.00446) \qquad\qquad\qquad (0.0039) \\
 & - 0.002203 \mathbf{DLPI_JAPAO}_{t-6} + 0.002514 \mathbf{DLPI_KOR}_{t-5} \\
 & (0.00062) \qquad\qquad\qquad (0.000882)
 \end{aligned}$$

Neste exemplo, temos 26 variáveis finalistas significantes no modelo final, com diferentes defasagens e variáveis representativas de praticamente todos os grupos, exceto mercado financeiro EUA e Balança comercial - demais países. Veja que temos a própria defasagem do PIB de 4 períodos (indicando uma certa sazonalidade), indicadores de produção industrial (bens de capital, indústrias extrativas, indústria geral, máquinas e equipamentos, alimentos e papel e celulose), utilização da capacidade instalada e horas pagas da indústria de São Paulo, taxa Selic, base monetária M0 e M1, despesas com execução financeira, exportações, exportações de veículos automotores - ônibus e máquinas agrícolas, e por fim a produção industrial de Japão e Coreia. A tabela 5 apresenta as variáveis mais significativas do modelo final.

6.3 Variáveis finais do modelo

Tabela 5: Principais variáveis no modelo final com seus respectivos coeficientes, defasagens e significância.

Coef.	Variável	Defasagem	Sinal	p-valor	Significância
0,5585	Utiliz. da capac. inst. indústria de SP	1 período	(-)	0,0666	Fraca (a 10%)
0,3378	Horas pagas na indústria de SP	4 períodos	(-)	0,0546	Fraca (a 10%)
0,3576	Horas pagas na indústria de SP	6 períodos	(+)	0,0691	Fraca (a 10%)
0,2458	Exportações FOB	2 períodos	(-)	0,0566	Fraca (a 10%)
0,2909	Exportações quantum	2 períodos	(+)	0,057	Fraca (a 10%)
0,2535	Prod. ind. de máquinas e equip.	contempor.	(+)	0,0291	Forte (a 5%)

Fonte: elaboração do próprio autor.

Os testes realizados para este modelo tiveram os seguintes resultados:

Teste	Resultado	p-valor
Normalidade	0.11689	0.9432
Heterocedasticidade	1.2838	0.1711
Autocorrelação	0.28251	0,9215

Veja que este resultado indica que os principais testes tiveram resultado esperado neste modelo, considerando a significância de 5%. Por outro lado, ao considerar variáveis *dummies* temos o seguinte resultado:

6.3 Variáveis finais do modelo

$$\begin{aligned}
 \text{DLPIB} = & - 0.332 \text{DLPIB}_{t-2} + 0.1096 \text{DLPIB}_{t-3} + 0.279 \text{DLPIB}_{t-4} \\
 & (0.00383) \quad (0.00511) \quad (0.00387) \\
 & + 0.001465 + 0.03209 \text{DLPI_BENS_NDUR}_{t-2} + 0.008888 \text{DLPI_BENS_NDUR}_{t-3} \\
 & (0.000268) \quad (0.00231) \quad (0.00251) \\
 & + 0.04201 \text{DLPI_BENS_NDUR}_{t-5} + 0.03157 \text{DLPI_EXTR_MIN}_{t-5} - 0.05772 \text{DLPI_EXTR_MIN}_{t-6} \\
 & (0.00328) \quad (0.00133) \quad (0.00117) \\
 & + 0.008109 \text{DLPI_FARMA}_{t-3} + 0.1089 \text{DLMAQ}_t + 0.06348 \text{DLMAQ_EQ}_{t-1} \\
 & (0.00102) \quad (0.00192) \quad (0.00182) \\
 & - 0.004691 \text{DLTRANSPO}_{t-5} - 0.05319 \text{DLTRANSPO}_{t-6} - 0.03802 \text{DLPI_QUIM}_t \\
 & (0.00144) \quad (0.00125) \quad (0.00215) \\
 & - 0.0417 \text{DLPI_QUIM}_{t-1} + 0.3202 \text{DLUTIL_CAP_INDSP}_t + 0.1307 \text{DLUTIL_CAP_INDSP}_{t-1} \\
 & (0.00189) \quad (0.00655) \quad (0.00714) \\
 & - 0.3299 \text{DLHR_PAG_IND}_t - 0.3305 \text{DLHR_PAG_IND}_{t-4} + 0.5463 \text{DLPES_OCUP_IND}_t \\
 & (0.00835) \quad (0.0068) \quad (0.0204) \\
 & - 0.131 \text{DLPES_OCUP_IND}_{t-3} - 0.001364 \text{DLSELIC}_t + 0.01448 \text{DLSELIC}_{t-6} \\
 & (0.016) \quad (0.000322) \quad (0.0003) \\
 & + 0.006468 \text{DLM0}_{t-3} + 0.007601 \text{DLM1}_{t-2} + 0.07307 \text{DLM1}_{t-4} \\
 & (0.000988) \quad (0.00132) \quad (0.000856) \\
 & + 0.05781 \text{DLDIV_INT_FED}_{t-1} + 0.01927 \text{DLEXEC_FIN_DESP}_{t-2} - 0.01756 \text{DLEXEC_FIN_DESP}_{t-4} \\
 & (0.000949) \quad (0.000397) \quad (0.000493) \\
 & - 0.1574 \text{DLIPCA}_{t-1} + 0.3889 \text{DLIPCA}_{t-2} - 0.04894 \text{DLIPCA}_{t-3} \\
 & (0.00596) \quad (0.00714) \quad (0.00191) \\
 & + 0.05556 \text{DLINCC}_{t-1} - 0.07518 \text{DLINCC}_{t-5} - 0.2902 \text{DLIGPDI}_{t-2} \\
 & (0.00551) \quad (0.00168) \quad (0.00694) \\
 & + 0.2663 \text{DLEXP_FOB}_{t-1} - 0.01624 \text{DLEXP_FOB}_{t-5} - 0.284 \text{DLEXP_QTUM}_{t-1} \\
 & (0.0043) \quad (0.00505) \quad (0.00455) \\
 & + 0.0216 \text{DLEXP_QTUM}_{t-5} - 0.02605 \text{DLEXP_MUND}_{t-5} - 0.01051 \text{DLEUA_JUR_NOT}_t \\
 & (0.00487) \quad (0.00342) \quad (0.00049) \\
 & - 1.008 \text{DLEUA_JUR_PRIME}_t - 0.1477 \text{DLEUA_JUR_PRIME}_{t-5} + 2.19 \text{DLIBOVESPA}_t \\
 & (0.0231) \quad (0.00454) \quad (0.0417) \\
 & + 0.005919 \text{DLNASDAQ}_{t-1} + 0.008037 \text{DLPETR_CONS}_{t-1} - 0.006365 \text{DLPETR_CONS}_{t-2} \\
 & (0.00213) \quad (0.00254) \quad (0.00207) \\
 & + 0.01014 \text{DLPETR_CONS}_{t-5} + 0.003444 \text{DLPETR_IMP}_{t-6} + 0.005195 \text{DLAUTO_LEV_PROD}_{t-6} \\
 & (0.00198) \quad (0.00024) \quad (0.00125) \\
 & - 0.002126 \text{DLAUTO_CAM_PROD}_{t-5} + 0.01044 \text{DLAUTO_CAM_PROD}_{t-6} - 0.01925 \text{DLAUTO_ONI_PROD}_{t-2} \\
 & (0.000479) \quad (0.000568) \quad (0.00034) \\
 & + 0.002486 \text{DLAUTO_PROD_TOT}_{t-4} + 0.009001 \text{DLAUTO_PROD_TOT}_{t-6} + 0.0005979 \text{DLAUTO_LEV_EXP}_{t-4} \\
 & (0.000763) \quad (0.00151) \quad (0.00026) \\
 & - 0.003744 \text{DLMAQ_EXP}_{t-1} + 0.01073 \text{DLMAQ_EXP}_{t-2} + 0.005848 \text{DLCOM_ENER_IND}_{t-1} \\
 & (0.000403) \quad (0.000377) \quad (0.00351) \\
 & - 0.0398 \text{DLCOM_ENER_IND}_{t-2} + 0.046 \text{DLCOM_ENER_COM}_{t-3} + 0.1362 \text{DLCOM_ENER_COM}_{t-5} \\
 & (0.00328) \quad (0.00296) \quad (0.00414) \\
 & - 0.03825 \text{DLCOM_ENER_RES}_{t-2} - 0.0981 \text{DLCOM_ENER_RES}_{t-5} - 0.1525 \text{DLCOM_ENER_OUT}_t \\
 & (0.00332) \quad (0.00438) \quad (0.00394) \\
 & - 0.1312 \text{DLCOM_ENER_OUT}_{t-4} + 0.0003445 \text{DLPI_ALEM}_{t-2} + 0.0003941 \text{DLPI_ALEM}_{t-5} \\
 & (0.0047) \quad (7.86e-005) \quad (5.66e-005) \\
 & + 0.002184 \text{DLPI_FRA}_t - 0.0002624 \text{DLPI_FRA}_{t-1} - 0.00259 \text{DLPI_JAPAO}_{t-5} \\
 & (3.81e-005) \quad (4.43e-005) \quad (5.91e-005) \\
 & - 0.009752 \text{I:1993(1)}_t - 0.00932 \text{I:1994(2)}_t + 0.01911 \text{I:1994(3)}_t \\
 & (0.00105) \quad (0.000891) \quad (0.00112) \\
 & - 0.03237 \text{I:1998(1)}_t + 0.004558 \text{I:1998(2)}_t + 0.01239 \text{I:2000(1)}_t \\
 & (0.000754) \quad (0.000704) \quad (0.000908) \\
 & + 0.01058 \text{I:2000(3)}_t - 0.007732 \text{I:2001(1)}_t - 0.006578 \text{I:2002(3)}_t \\
 & (0.000806) \quad (0.000757) \quad (0.000961) \\
 & + 0.002845 \text{I:2003(3)}_t - 0.01455 \text{I:2004(2)}_t - 0.009122 \text{I:2006(2)}_t \\
 & (0.000793) \quad (0.000823) \quad (0.000812) \\
 & + 0.03906 \text{I:2006(3)}_t + 0.02979 \text{I:2007(3)}_t + 0.006539 \text{I:2010(2)}_t \\
 & (0.000839) \quad (0.00076) \quad (0.00104) \\
 & + 0.00883 \text{I:2010(3)}_t \\
 & (0.000917)
 \end{aligned}$$

Neste caso tivemos 72 variáveis significativas no modelo final e 16 variáveis *dummies*, também distribuídas em todas as defasagens. Além das próprias defasagens do PIB, grande parte de variáveis de produção industrial (12), capacidade da indústria de São Paulo (6), taxa SELIC

6.3 Variáveis finais do modelo

Tabela 6: Principais variáveis no modelo final com seus respectivos coeficientes, defasagens e significância.

(2), base monetária (3), dívida federal (3), IPCA (3), INCC (2), IGPDÍ (1), exportações (5), taxa de juros americana (3), Ibovespa (1), NASDAQ (1), consumo e importação de petróleo (4), produção de veículos automotores (9), consumo de energia (8), e produção industrial dos países desenvolvidos (5). Além disto, note que tivemos a presença de 16 variáveis *dummies*: I:1993(1) - plano cruzeiro real, I:1994(2,3) - plano real, I:1998(1,2) - reeleição do Fernando Henrique Cardoso, I:2000(1,3), I:2001(1), I:2002(3) - eleição de Luis Inácio Lula da Silva para presidente, I:2003(3), I:2004(2), I:2006(2,3), I:2007 (3), I:2010(2,3). A tabela 6 apresenta as variáveis mais significativas do modelo final.

Coef.	Variável	Defasagem	Sinal	p-valor	Significância
2,19	IBOVESPA	contempor.	(+)	0,0417	Forte (a 5%)
1,008	Tx. juros prime EUA	contempor.	(-)	0,0231	Forte (a 5%)
0,5463	Pessoal ocupado na ind. de SP	contempor.	(+)	0,0204	Forte (a 5%)
0,3889	IPCA	2 períodos	(+)	0,00714	Forte (a 5%)
0,332	PIB	2 períodos	(+)	0,00383	Forte (a 5%)
0,3305	Horas pagas ind. de SP	4 períodos	(-)	0,0068	Forte (a 5%)
0,3299	Horas pagas ind. de SP	contempor.	(-)	0,00835	Forte (a 5%)
0,3202	Utiliz. da capac. inst. indústria de SP	contempor.	(+)	0,00655	Forte (a 5%)
0,2902	IGPDÍ	2 períodos	(-)	0,00694	Forte (a 5%)
0,284	Exportações Quantum	1 período	(-)	0,00455	Forte (a 5%)
0,279	PIB	4 períodos	(+)	0,00387	Forte (a 5%)
0,2663	Exportações FOB	1 período	(+)	0,0043	Forte (a 5%)

Fonte: elaboração do próprio autor.

Os testes realizados para este modelo tiveram os seguintes resultados:

6.4 Sugestão para pesquisas

Teste	Resultado	p-valor
Normalidade	6,1273	0,0467
Heterocedasticidade	sem obs.	
Autocorrelação	1.7013	0.1744

Veja que este resultado indica que os principais testes tiveram resultado esperado neste modelo, considerando a significância de 5%.

6.4 Sugestão para pesquisas

Como próximos passos deste trabalho, indicamos a reestimação dos modelos considerando a nova metodologia de cálculo do PIB lançado pelo IBGE em meados de Março/2015. Não se espera grandes alterações nos resultados apresentados, porém, a atualização pode ser de valiosa importância para comparação e avaliação de impactos com relação aos resultados deste trabalho.

Além disto, sempre podemos aumentar o conjunto de índices e variáveis explicativas utilizadas, levando-se em conta o período necessário e a disponibilidade das mesmas pelos órgãos competentes respectivos. Outra extensão possível deste trabalho é realizar os mesmos testes e análises para países com ambiente econômico semelhante ao Brasil, como, por exemplo, países da América Latina, para posterior comparação com o caso brasileiro.

Outra sugestão é a utilização de *dummies de step* e iteração entre *step* regressões para permitir estruturas variantes no tempo. Podemos também utilizar modelos fatoriais, modelos FAVAR (Benanke) e ainda comparar os resultados destes modelos com os obtidos em (Hendry, Castle e Kitov, (2013)).

7 Conclusão

Ao longo deste trabalho apresentamos o resultado do exercício de previsão de alguns modelos de previsões do PIB brasileiro utilizando técnicas de *nowcasting*, com base em uma seleção inicial de variáveis que foram relevantes em algum momento da economia brasileira na previsão da atividade econômica. Partimos de 72 variáveis explicativas, além das próprias defasagens do PIB, para tentar estimar um modelo que fosse estatisticamente mais robusto que um modelo AR utilizado como *benchmark*, através de um algoritmo de seleção de variáveis.

Como primeiro resultado podemos concluir que os modelos propostos foram, no mínimo, iguais ou melhores que o modelo utilizado como benchmark em todos os casos sugerindo que técnicas de *nowcasting* tem muita informação relevante. Mesmo considerando uma economia em desenvolvimento, estes resultados confirmaram os resultados obtidos em (Castle, Hendry e Kitov, (2013)) para uma economia mais estável.

Outro resultado relevante deste trabalho foi uma característica a parte dos dados brasileiros. A utilização de variáveis *dummies* não foi decisiva para definir a melhor eficácia das previsões, enquanto que a escolha da significância utilizada na estimação e no algoritmo de seleção de variáveis mostrou-se relevante, e um importante fator de diferenciação e eficácia de modelos de previsão e *nowcast* no caso brasileiro. A utilização deste parâmetro em 1% mostrou desempenho acima do valor de 0,1%. Este resultado não foi observado, por exemplo, no trabalho de e (Castle, Hendry e Kitov, (2013)), e entendemos tratar-se de característica intrínseca dos dados obtidos da economia brasileira.

Por último, tivemos um resultado ambíguo no que tange o papel dos ciclos econômicos no desempenho preditivo dos modelos. Para um dos modelos considerados (NC3), concluímos que utilizando a significância de seleção de modelos em 1%, e com uso de variáveis *dummies*, por exemplo, o desempenho médio do modelo é superior nos períodos de recessão se comparado aos períodos de expansão.

A Apêndice

A.1 Lista de indicadores e suas defasagens utilizados no modelo

Na figura 8, apresentamos a lista completa dos indicadores utilizados neste trabalho.

Tipo	Grupo	Variável	Nome Base	Período	Início	Divulgação	Defasagem	Origem
Real	Produção Industrial (16)	Índice de Produção Industrial - Bens de consumo	PI_BENS_CONS	Mensal	1975	1a Quinz	2 meses	IBGE
		Índice de Produção Industrial - Bens Intermediários	PI_BENS_INT					
		Índice de Produção Industrial - Bens de Capital	PI_BENS_CAP					
		Produção industrial - bens de consumo semi e não duráveis	PI_BENS_NDUR					
		Índice de Produção Industrial - Bens de Consumo duráveis	PI_BENS_DUR					
		Produção industrial - indústrias extrativas - quantum	PI_EXTR_MIN					
		Produção industrial - produtos farmoquímicos e farmacêuticos	PI_FARMA					
		Produção industrial - indústria geral	PI_IND_GER					
		Produção industrial - máquinas e equipamentos	MAQ_EQ					
		Produção industrial - metalurgia	PL_METALU					
		Produção industrial - outros equipamentos de transporte	TRANSP					
		Produção industrial - produtos alimentícios	PI_ALIMEN					
		Produção industrial - papel e papelão	PI_PAPEL					
		Produção industrial - produtos de borracha e material plástico	PI_PLAS_BORR					
Produção industrial - química	PI_QUIM							
Produção industrial - produtos têxteis	PI_TEXT							
		PIB do Brasil	PIB	Trimestral				
Real	Produção Auto	Produção Automóveis	AUTO_PROD	Mensal	1957	1a Quinz	1 mês	Anfavea
		Produção comerciais leves	AUTO_LEV_PROD					
		Produção caminhões	AUTO_CAM_PROD					
		produção ônibus	AUTO_ONI_PROD					
		Produção - autoveículos - montados	AUTO_PROD_TOT					
		Produção - Máquinas Agrícolas e Rodoviárias	MAQ_PROD					
Real	Mercado de trabalho	Utilização da capacidade instalada - indústria - SP	UTIL_CAP_INDSP	Mensal	1975	1a Quinz	1 mês	FIESP
		Horas pagas - indústria	HR_PAG_IND					
		Índice de horas trabalhadas na indústria de SP	HR TRAB_IND					
		Pessoa ocupada - indústria - SP	PES_OCUP_IND					
Real	Consumo - Energia Elétrica	Consumo energia elétrica - indústria	COM_ENER_IND	Mensal	1976	1a Quinz	2 meses	Eletrobras
		Consumo energia elétrica - comércio	COM_ENER_COM					
		Consumo energia elétrica - residencia	COM_ENER_RES					
		Consumo energia elétrica - outros	COM_ENER_OUT					
		Consumo energia elétrica - total	COM_ENER_TOT					
Financeira	Mercado financeiro	Taxa de juros - Over / Selic	SELIC		1974	1a Quinz	-	BCB
		MO - base monetária - fim período	MO		1946	1a Quinz	1 mês	BCB
		M1 - fim período	M1		1946	1a Quinz	1 mês	BCB
		Dívida mobiliária interna federal - títulos fora do BC - fim período	DIV_INT_FED		1970	2a Quinz	1 mês	BCB
		Execução financeira - despesas - total	EXEC_FIN_DESP		1973	1a Quinz	2 meses	Min Fazenda
		Execução financeira - receitas - total	EXEC_FIN_REC		1975	1a Quinz	2 meses	Min Fazenda
		IPCA	IPCA		1980	1a Quinz	1 mês	IBGE
		Ibovespa	IBOVESPA		1968	1a Quinz	-	Anbima
		Inflação - INCC	INCC		1944	2a Quinz	1 mês	FGV
		Inflação - IGP-DI	IGPDI		1944	1a Quinz	1 mês	FGV
		Financeira	Balança Comercial Brasil	Exportações (FOB)	EXP_FOB		1977	
Exportações - quantum - índice (média 2006 = 100)	EXP_QTUM				1977		2 meses	Funcex
Importações	IMP_QTUM				1978		2 meses	Funcex
Importações - (FOB)	IMP_FOB				1980		1 mês	Min Desenv
Taxa de câmbio - R\$ / US\$ - comercial - compra - fim período	TX_CAMBIO				1953		1 mês	BCB
Reservas internacionais - liquidez internacional	RES_INTER				1970	1a Quinz	1 mês	BCB
Exportação Automóveis	AUTO_EXP						1 mês	Anfavea
Exportação comerciais leves	AUTO_LEV_EXP						1 mês	Anfavea
Exportação caminhões	AUTO_CAM_EXP				1957		1 mês	Anfavea
Exportação ônibus	AUTO_ONI_EXP						1 mês	Anfavea
Exportação - autoveículos - montados	AUTO_EXP_TOT						1 mês	Anfavea
Exportação - Máquinas Agrícolas e Rodoviárias	MAQ_EXP				1 mês	Anfavea		
Financeira	Balança Comercial - Demais Países	Mundial - Exportações	EXP_MUND	Mensal	1957	2a Quinz	2 meses	FMI
		Mundial - Importações	IMP_MUND					
		Países industrializados - importações	IMP_PAIS_IND					
		Países industrializados - exportações	EXP_PAIS_IND					
Financeira	Mercado Financeiro - EUA	Estados Unidos - taxa de juros - fundos federais	EUA JUR_FED		1957	2a Quinz	2 meses	FMI
		Estados Unidos - taxa de juros - Notas do Tesouro	EUA JUR_NOT		1957	2a Quinz	2 meses	FMI
		Estados Unidos - taxa de juros - Notas do Tesouro - maturidade 3	EUA JUR_NOT3M		1974	2a Quinz	2 meses	FMI
		Estados Unidos - taxa de juros - Prime	EUA JUR_PRIME		1957	2a Quinz	2 meses	FMI
		Estados Unidos - índice de ações: NASDAQ - fechamento	NASDAQ		1971	1a Quinz	1 mês	Nasdaq
Commodities	Preços e Produção	Commodities - petróleo - cotação internacional	PETR_COTAC		1957	2a Quinz	2 meses	FMI
		Consumo aparente - derivados de petróleo - média - qde./dia	PETR_CONS		1979	2a Quinz	2 meses	ANP
		Importações - petróleo bruto - (FOB)	PETR_IMP		1984	1a Quinz	1 mês	Min Desenv
		Produção - petróleo - média - qde./dia	PETR_PROD		1979	2a Quinz	2 meses	ANP
Real	Produção Industrial - Resto do Mundo	Produção Industrial - EUA	PI_EUA			1a Quinz	2 meses	
		Produção Industrial - Alemanha	PI_ALEM			1a Quinz	3 meses	
		Produção Industrial - França	PI_FRA		1980	1a Quinz	3 meses	FMI
		Produção Industrial - Japão	PI_JAPAO			1a Quinz	3 meses	
		Produção Industrial - Korea	PI_KOR			1a Quinz	1 meses	

Figura 8: Lista completa dos indicadores utilizados neste trabalho, categorizadas de acordo com seu Tipo, Grupo, Periodicidade, início da série, período de divulgação dentro do mês, defasagem da informação e órgão responsável.

A.2 Ciclos Econômicos

Na figura 9, ilustramos o agrupamento das variáveis utilizadas no trabalho, com os respectivos períodos de publicação e defasagens. A primeira linha refere-se à quinzena de publicação da variável no mês, e a segunda é a defasagem da informação contida neste *release*. Na figura é mostrado um trimestre qualquer, e os 2 meses anteriores, e D1, D2 e D3 são as respectivas publicações mensais dentro do trimestre.

INDICADORES	Qt Vars	Trimestre 1									
		mês - 2		mês - 1		Mês 1		Mês 2		Mês 3	
		Q1	Q2	Q1	Q2	Q1	Q2	Q1	Q2	Q1	Q2
Produção Industrial e Consumo de energia elétrica	17										
Divulgação						D1		D2		D3	
Referência		D1		D2		D3					
Produção Auto e Mercado de Trabalho - Indústria - Demais	9										
Divulgação						D1		D2		D3	
Referência				D1		D2		D3			
Mercado de Trabalho - Capac. Instalada ind. - SP	1										
Divulgação							D1		D2		D3
Referência				D1		D2		D3			
Tx Juros - Selic, Ibovespa, Nasdaq	1										
Divulgação						D1	D2	D3			
Referência						D1	D2	D3			
Base Monetária M0 / M1, IPCA e IGP-DI e Exp. e Importações (FOB)	4										
Divulgação						D1		D2		D3	
Referência				D1		D2		D3			
Dívida mobiliária interna federal e INCC	2										
Divulgação							D1		D2		D3
Referência				D1		D2		D3			
Exportações e importações quantum e Execução Financeira	3										
Divulgação						D1		D2		D3	
Referência		D1		D2		D3					
Tx Cambio / Reservas Internac. E Exportações AUTO	2										
Divulgação						D1		D2		D3	
Referência				D1		D2		D3			
Mundial - Exp/Imp e Exp Países Ind, EUA - Taxas de Juros	8										
Divulgação							D1		D2		D3
Referência		D1		D2		D3					
Petróleo - Cotação Internacional e consumo	2										
Divulgação							D1		D2		D3
Referência		D1		D2		D3					
Importações - petróleo bruto	1										
Divulgação						D1		D2		D3	
Referência				D1		D2		D3			
Produção petróleo e Produção Industrial - Resto do Mundo	6										
Divulgação							D1		D2		D3
Referência		D1		D2		D3					
TOTAL	72										

Figura 9: Defasagens das variáveis explicativas utilizadas nos modelos de *forecasting* e *now-casting*.

A.2 Ciclos Econômicos

As tabelas a seguir utilizam ($\alpha = 1\%$) à esquerda e ($\alpha = 0,1\%$) à direita. Apresentando os resultados para o modelo NC2 temos:

A.2 Ciclos Econômicos

Ciclo econômico	sem IIS	com IIS	%	Ciclo econômico	sem IIS	com IIS	%
Expansão	7,51E-05	3,57E-05	48%	Expansão	1,48E-04	7,20E-04	486%
Recessão	1,07E-04	3,25E-05	30%	Recessão	7,79E-04	2,75E-03	353%
Relação	143%	91%		Relação	526%	381%	

Apresentando os resultados para o modelo NC1 temos:

Ciclo econômico	sem IIS	com IIS	%	Ciclo econômico	sem IIS	com IIS	%
Expansão	1,17E-03	1,11E-04	9%	Expansão	1,95E-04	1,62E-03	831%
Recessão	3,15E-04	1,49E-04	47%	Recessão	3,31E-04	1,06E-03	319%
Relação	27%	134%		Relação	170%	65%	

Apresentando os resultados para o modelo F3 temos:

Ciclo econômico	sem IIS	com IIS	%	Ciclo econômico	sem IIS	com IIS	%
Expansão	5,25E-03	1,68E-04	3%	Expansão	2,95E-03	2,98E-04	10%
Recessão	4,04E-04	3,78E-04	93%	Recessão	5,71E-04	4,77E-04	84%
Relação	8%	225%		Relação	19%	160%	

Apresentando os resultados para o modelo F2 temos:

Ciclo econômico	sem IIS	com IIS	%	Ciclo econômico	sem IIS	com IIS	%
Expansão	1,06E-04	1,43E-04	136%	Expansão	2,60E-04	1,63E-04	62%
Recessão	2,79E-04	2,79E-04	100%	Recessão	2,99E-04	2,35E-04	79%
Relação	264%	195%		Relação	115%	145%	

Apresentando os resultados para o modelo F1 temos:

Ciclo econômico	sem IIS	com IIS	%	Ciclo econômico	sem IIS	com IIS	%
Expansão	1,34E-04	9,09E-05	68%	Expansão	3,79E-04	5,26E-04	139%
Recessão	1,46E-04	6,99E-05	48%	Recessão	4,41E-04	3,78E-04	86%
Relação	109%	77%		Relação	116%	72%	

B Bibliografia

- AASTVEIT, K. A. e TROVIK, T. G. (2007). Nowcasting norwegian GDP: the role of asset prices in a small open economy. *Norges Bank - Working Paper*.
- BANBURA M., GIANNONE D. e REICHLIN L. (2010). Nowcasting. *Working Paper Series No 1275 - European Central Bank*.
- BANBURA M., GIANNONE D., MODUGNO M. e REICHLIN L. (2013). Now-casting and the real-time data flow. *Working Paper Series no 1564 - European Central Bank*.
- BRAGOLI D., METELLI L. e MODUGNO M. (2014). The importance of updating: evidence from a brazilian nowcasting model. *Finance and Economics Discussion Series - Federal Reserve Board, Washington, D.C.*
- CASTLE J. L., DOORNIK J. A. e HENDRY, D. F. (2012). Model selection when there are multiple breaks. *Journal of Econometrics - Volume 169 - Numero 2*.
- CASTLE J. L., HENDRY D. F. e KITOV O. I. (2013). Forecasting and Nowcasting macroeconomic variables: a methodological overview. *University of Oxford - Department Of Economics*.
- CHAUVET, M. (2000). A monthly indicator of brazilian GDP.
- CHAUVET M., LIMA, E. C. R. e VASQUEZ B. (2002). Forecasting brazilian output in real time in the presence of breaks: a comparison of linear and nonlinear models. *IPEA - Texto para discussao No 911*.
- D'AGOSTINO A., GIANNONE D., LENZA M. e MODUGNO M. (2014). Nowcasting business cycles: a bayesian approach to dynamic heterogeneous factor models. *Work in progress*.
- ERICSSON, N. R. (2012). How biased are U.S. government forecasts of the federal debt. *In: Federal Reserve Board and George Washington University*.
- FERRARA, L. e MARSILLI, C. (2014). Nowcasting global economic growth: a factor-augmented mixed-frequency approach. *Document de travail no 515 - Banque de France*.

- FERRARA L., GUEGAN D. e RAKOTOMAROLAHY P. (2010). GDP nowcasting with ragged-edge data: a semi-parametric modeling. *Journal of Forecasting, Wiley-Blackwell, 2010, 29 (1-2)*.
- FERREIRA, R. T. (2005). *Forecasting quarterly brazilian GDP growth rate with linear and nonlinear diffusion index models*. Ph.D. thesis, Universidade Federal do Ceara.
- GIANNONE D., REICHLIN L. e SMALL D. H. (2006). Nowcasting GDP and inflation. the real-time informational content of macroeconomic data releases. *Working Paper Series no 633 - European Central Bank*.
- LIU P., MATHESON T. e ROMEU R. (2011). Real-time forecasts of economic activity for latin american economies. *IMF Working Paper - Western Hemisphere and Research Departments*.
- MODUGNO M., SOYBILGEN B. e YAZGAN E. (2015). Nowcasting Turkish GDP and News Decomposition. *Preliminary paper - Work in progress*.
- VIEIRA, H. P. e PEREIRA, P. L. V. (2014). Um estudo sobre os ciclos de negocios brasileiro (1900-2012). *Working Paper Series - CEQEF No 14 - FGV*.
- YIU, M. S. e CHOW, K. K. (2011). Nowcasting chinese GDP: information content of economic and financial data. *Hong Kong Institute for Monetary Research - Volume 3*.