

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



MONOGRAFIA DE FINAL DE CURSO

NOWCASTING DE PIB NO BRASIL

Aluno: Rafael Lima da Fonseca

Matrícula: 1311042

Orientador: Marcelo Medeiros

Junho de 2017

PUC-RIO

2017.1



MONOGRAFIA DE FINAL DE CURSO

NOWCASTING DE PIB NO BRASIL

Aluno: Rafael Lima da Fonseca

Matrícula: 1311042

Orientador: Marcelo Medeiros

Declaro que o presente trabalho é de minha autoria e que não recorri para realiza-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor

Junho de 2017

PUC-RIO

2017.1

TERMO DE RESPONSABILIDADE

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor

Sumário

Introdução.....	5
Revisão de literatura.....	8
Dados.....	12
Metodologia.....	15
Resultados.....	18
Conclusão.....	25
Referências Bibliográficas.....	26

Lista de tabelas e gráficos

TABELAS

Tabela 1.....	14
Tabela 2.....	21
Tabela 3.....	21
Tabela 4.....	22
Tabela 5.....	22

GRÁFICOS

Gráfico 1.....	18
Gráfico 2.....	19
Gráfico 3.....	20
Gráfico 4.....	23
Gráfico 5.....	23

I - INTRODUÇÃO

Dentro da teoria macroeconômica moderna, política monetária é vista como a principal arma a ser usada para combater acelerações inflacionárias e diminuições no ritmo de atividade da economia. Dessa forma, é fundamental que a autoridade monetária saiba reconhecer em que momento do ciclo econômico a economia se encontra. Além disso, como a política monetária age com *lags* sobre a economia, é necessário também que a autoridade monetária tenha uma razoável noção do futuro próximo da economia.

Isso se torna um desafio grande quando a maioria dos dados macroeconômicos utilizados como termômetro da economia são divulgados meses depois do seu período de referência. Por exemplo, o dado de desemprego que temos para o Brasil sai com aproximadamente um mês de atraso (ao final de junho, temos acesso ao número de maio).

Outro problema com o qual a autoridade monetária se depara é o elevado número de séries macroeconômicas que poderiam ser usadas para medir a performance da economia. Isso se torna especialmente problemático quando esses diversos indicadores apontam para direções diferentes. Muitas vezes, poderia ser mais fácil olhar para alguns poucos indicadores que sintetizam toda essa informação.

Um dos indicadores que poderia servir essa função é o Produto Interno Bruto (PIB). Como o PIB é um cálculo de tudo o que é produzido pela economia, ele pode servir como importante síntese dos outros indicadores de atividade econômica. Além disso, o PIB permite uma análise de quais são os setores da economia que estão melhor ou pior. Isso pode ser importante para a política monetária já que uma alteração na taxa de juros pode afetar esses grupos de maneiras diferentes e em velocidades distintas.

Porém, o PIB é um dos dados que mais sofre do problema de falta de contemporaneidade. Alguns bancos centrais, reconhecendo esse problema criaram estatísticas que servem como *proxy* para o PIB (é o caso do IBC-Br no Brasil, por exemplo). Apesar de suas divulgações serem mais oportunas que a do PIB, esses índices ainda sofrem com divulgações com certo grau de defasagem.

Foi exatamente pensando nessa situação que a técnica de *nowcasting* foi criada. Ela permite que se tenha uma estimativa para o PIB em um futuro próximo (mais especificamente, para o trimestre imediatamente posterior ao último trimestre que se tem

dados oficiais para o PIB) considerando uma grande quantidade de informações relevantes à medida que elas se tornam disponíveis e levando em conta sua contemporaneidade.

Diversos bancos centrais ao redor do mundo já utilizam modelos de *nowcasting* atualizados regularmente. Podemos destacar modelos usados pelo BCE (Banco Central Europeu); FRB Atlanta; FRB Nova Iorque (ambos bancos regionais ligados ao sistema do FED, Federal Reserve, dos EUA); entre outros.

No curso de seu uso em diversos contextos e situações, esses modelos mostraram que tem uma boa performance preditiva, apesar de normalmente não serem tão precisos, em média, quanto a mediana das expectativas de mercado (usualmente, o melhor previsor para uma série de estatísticas econômicas). Porém, esses modelos, diferentemente das expectativas de mercado, são muito transparentes, permitindo enxergar os fatores que levaram a mudanças nas estimativas.

O objetivo desta monografia foi construir um modelo de *nowcasting* para a economia brasileira se baseando na metodologia empregada em Giannone, Reichlin e Small (2008). Esse modelo específico é utilizado atualmente pelo FRB Nova Iorque para realizar o seu *nowcasting* de PIB para os EUA. Além disso, esse método específico já foi adaptado para diferentes contextos e com razoável grau de sucesso. A grande vantagem de usar essa metodologia (quando comparada, por exemplo, à técnica empregada pelo FRB Atlanta¹) é que sua flexibilidade permite que sua aplicação seja feita praticamente sem adaptações específicas ao país/região que está sendo aplicada. Dessa forma, todas as características desejáveis de eficiência da técnica² podem ser mantidas no modelo final.

Os resultados obtidos para o caso brasileiro atestam para a qualidade e flexibilidade do modelo. Usando a metodologia de Giannone, Reichlin e Small (2008) sem nenhuma adaptação específica para o caso brasileiro, o modelo conseguiu ter uma performance *out of sample* ligeiramente melhor que modelos estatísticos ingênuos (que normalmente costumam ser muito difíceis de serem batidos) e chegar próximo da previsão da pesquisa FOCUS.

¹ Descrita em Higgins (2014)

² Descrita em Doz, Giannone e Reichlin (2009)

O resto do texto está organizado da maneira a seguir. A seção 2 contém uma revisão da literatura relevante. A seção 3 descreve em mais detalhes os dados utilizados para a estimação do modelo. A seção 4 descreve a metodologia empregada para a confecção do modelo. A seção 5 descreve os resultados do modelo e os compara com outras estimativas. Por fim, a seção 6 conclui.

II – REVISÃO DE LITERATURA

Apesar de ser um método estatístico relativamente recente, a literatura de *nowcasting* já é vasta, com aplicações da estratégia para vários contextos diferentes. É importante entender os diversos resultados e as suas especificidades metodológicas para a confecção de um modelo adequado para o caso brasileiro.

Primeiramente, é importante compreender exatamente o que é *nowcasting*. Banbura, Giannone e Reichlin (2011) vão definir a técnica como “a previsão do presente, do futuro próximo ou do passado recente”³. Aqui é fundamental entender o que os autores chamam de “previsão do presente (...) ou do passado recente” já que parece contra intuitivo a previsão de algo que já passou. Estatísticas econômicas, de maneira geral, são divulgadas com alguma defasagem do momento que elas retratam. Por exemplo, o PIB brasileiro para o terceiro trimestre de 2016 foi divulgado no dia 29 de novembro de 2016, ou seja, na metade do quarto trimestre. Assim, se existisse no início de novembro uma previsão para um PIB do terceiro trimestre, teríamos uma informação relevante que seria uma “previsão do passado recente”, da forma que os autores estão se referindo.

Banburra, Giannone e Reichlin vão argumentar que a técnica é melhor aproveitada se utilizada em variáveis que são publicadas com pouca frequência no ano e com largas defasagens (por exemplo, o PIB que é publicado apenas quatro vezes ao ano com uma defasagem considerável como demonstrado pelo caso brasileiro citado acima). Isso permite o uso de séries com publicação mais frequente e com menos defasagens para prever a evolução da variável em questão.

Na visão de Banbura Giannone e Reichlin, a técnica de *Nowcasting* não é uma simples agregação de dados, a partir de um conjunto de equações (tipicamente chamadas de *bridge equations*), para produzir uma estimativa oportuna de um indicador. Um modelo de *nowcasting* irá atualizar sua previsão à medida que novas informações são disponibilizadas. A partir dessa série de previsões, é possível entender quais são as variáveis que tem maior impacto na estimativa. Isso permite uma interpretação mais qualitativa das estimativas. Por exemplo, se a previsão para o PIB foi reduzida por conta de um dado de produção industrial fraco, podemos ir um passo além e especular que

³ “The problem of predicting the presente, the very near future and the very recent past is labelled as *nowcasting*” no original

provavelmente o componente de produção industrial no PIB deve vir fraco também, algo mais difícil de se fazer caso a previsão não reagisse à novas informações.

Existem diversas maneiras de se atingir esse mesmo objetivo. Uma delas, é a proposta por Giannone, Reichlin e Small (2008). Os autores pretendem realizar um *nowcasting* para o PIB dos EUA. Para isso eles contam com uma base de dados de aproximadamente 200 variáveis. Devido à grande quantidade de variáveis, uma estimação do modelo completo levaria à muita incerteza em relação aos parâmetros do modelo e, conseqüentemente, uma performance ruim do mesmo. Para solucionar esse problema, os autores usam a técnica de previsão a partir de um modelo de fatores comuns, estimados a partir de componentes principais⁴. Porém, dentro da base de dados compilada pelos autores, existem dados com a última observação em períodos de tempo diferentes (por exemplo, no início de um mês, digamos dezembro, já existe uma observação da *ISM PMI Survey*, uma importante pesquisa sobre vendas no setor manufatureiro, para novembro, porém o último dado de produção industrial é referente a outubro). Isso inviabiliza a estimação de fatores pela maneira usual (já que para a estimação por componentes principais requer um painel balanceado). Para solucionar esse problema, os autores aplicam uma técnica de estimação em dois estágios usado em Giannone, Reichlin e Sala (2004). Primeiro os fatores comuns são estimados normalmente usando-se o maior painel balanceado possível. Esses fatores são usados para estimar parâmetros que, posteriormente, são utilizados pelo filtro de Kalman para estimar os fatores comuns usando a base de dados completa, ou seja, incluindo momentos no tempo onde nem todas as séries tem observações. Com os fatores em mãos, os autores obtêm o PIB como uma função linear dos fatores comuns esperados. A cada nova divulgação de dado, as estimativas para os fatores comuns são atualizadas e, conseqüentemente, a estimativa para o PIB também.

Os resultados obtidos por Giannone, Reichlin e Small sugerem que o modelo tem boa performance na previsão *out of sample* para a próxima divulgação do PIB, sendo melhor que uma previsão de crescimento constante ou a previsão da *Survey of Professional Forecasters* (uma pesquisa trimestral em que especialistas são perguntados suas previsões para diversas variáveis macroeconômicas, entre elas, o crescimento). Para previsões com um horizonte de tempo maior (um, dois, três ou quatro trimestres a frente)

⁴ Stock e Watson (2011) fazem uma revisão de literatura relacionada a essa técnica.

a previsão de crescimento constante é melhor (o modelo de *nowcasting* porém continua sendo mais preciso que a *Survey of Professional Forecasters*). Dessa forma, os autores sugerem que as vantagens trazidas pela técnica de *nowcasting* são apenas observadas quanto se faz a previsão para o curtíssimo prazo. Além disso, os autores mostram que, como o esperado, a previsão fica mais precisa a medida que mais informações são incluídas no modelo, sendo a última estimativa antes da divulgação do PIB a que apresenta menor erro, em média, da série de estimativas criadas.

Uma outra forma de se desenhar um modelo de *nowcasting* é a escolhida por Higgins (2014) na confecção do modelo GDPNow usado pelo FRB Atlanta. Diferentemente do modelo de Giannone, Reichlin e Small, além de gerar uma previsão para o crescimento do PIB no trimestre, o modelo tem estimativas para cada um dos componentes do PIB. Para conseguir isso, o autor primeiro estima um BVAR trimestral com os principais componentes do PIB e gera uma previsão para cada componente para o próximo trimestre. Depois, ele divide cada componente em seus subcomponentes. Para cada subcomponente, ele escolhe uma série mensal relacionada para fazer a previsão do mesmo. A série mensal é transformada em trimestral ao se fazer uma média dos três meses relevantes. Para dados em que todos os meses ainda não estão disponíveis, é feita uma previsão usando um fator comum estimado, usando a mesma técnica de Giannone, Reichlin e Small, a partir de uma base de dados de 124 séries mensais. As previsões dos subcomponentes são então agregadas, criando uma estimativa para cada um dos componentes. É então feita uma média entre as previsões dos componentes a partir do BVAR e a partir da previsão dos subcomponentes, criando assim a estimativa final de cada componente. Por fim, os componentes são agregados, com os mesmos pesos usados na divulgação do PIB, gerando assim, a previsão final de crescimento. A medida que novas informações são divulgadas, tanto o fator é atualizado, quanto, possivelmente as séries escolhidas para a previsão dos subcomponentes. Dessa forma, a previsão para o PIB é atualizada ao longo do tempo.

Higgins mostra que o GDPNow tem uma performance *out of sample* melhor que outros possíveis modelos estatísticos para prever o crescimento. Porém, o autor acha que, provavelmente, o consenso das expectativas de mercado no longo prazo é mais preciso que a previsão do modelo. Em particular, estimativas baseadas em especialistas parecem ter uma performance melhor especialmente nos componentes de gastos de governo e variação de estoque. Posteriormente, Higgins (2016) vai mostrar que, não só o GDPNow

tem tido uma performance relativamente similar ao consenso de mercado (na realidade, a performance do modelo tem sido ligeiramente melhor) como conseguiu isso ao mesmo tempo em que sua previsão era bastante diferente do consenso, em média (de maneira geral, previsões individuais que são, em média, muito diferentes do consenso tendem a ser piores). Assim como o modelo de Giannone, Reichlin e Small, a estimativa fica mais precisa à medida que o tempo passa. Porém, é interessante notar que nas quatro últimas semanas antes da divulgação do dado oficial de PIB, as estimativas do modelo tem erro, média, muito similares.

Os dois exemplos que foram focados são de modelos para a previsão de PIB nos EUA. Porém essa técnica já foi aplicada também em outros contextos. Angelini et al. (2011) usam um modelo similar de Giannone, Reichlin e Small para fazer a previsão para o PIB da zona do euro enquanto Runstler et al. (2009) usam a técnica para fazer a previsão para cada um dos países da zona do euro. *Nowcasting* também foi usado para previsões do PIB da Nova Zelândia (Matheson 2010), Noruega (Aastveit e Trovik, 2008), China (Yiu e Chow, 2011), entre outros. Já Modugno (2011) usa um modelo de *Nowcasting* para fazer previsões para a inflação (um dado mensal), a partir de séries semanais e diárias⁵, para a zona do euro e os EUA.

⁵ Foroni e Marcellino (2013) fazem uma revisão da literatura sobre modelos econométricos com dados de diferentes frequências.

III - DADOS

O primeiro passo para se realizar o exercício de *nowcasting* é a criação da base de dados a ser utilizada. A base de dados montada é composta de 68 séries temporais mensais mais a série trimestral do PIB. A série encadeada do PIB com a metodologia atual do IBGE começa em 1996. Para o uso da técnica, é necessário que cada uma das séries individuais seja tão longa quanto a série de PIB. Dessa forma, séries que tenham seu início posterior a 1996 não foram incluídas na base de dados. Isso inclui dados considerados relevantes, como o IBC-Br e o desemprego. Porém, caso outras séries, que estejam na base de dados, sejam suficientes para resumir a informação contida nas séries que não foram incluídas, suas ausências não irão impactar o resultado final. Naturalmente, para séries que são mais longas que o PIB (como por exemplo, o IPCA), apenas o pedaço a partir de 1996 é utilizado.

Era fundamental ter duas informações para cada uma das séries. A primeira delas é a série histórica dos dados escolhidos. A segunda informação importante era a data de divulgação de cada uma das séries.

Primeiramente sobre a compilação das séries históricas. Muitas séries temporais mensais sofrem constantes revisões. Existem diversos motivos para que essas revisões ocorram (séries podem sofrer mudanças metodológicas ou o instituto que publica a série pode ter acesso a dados melhores para a confecção da série, por exemplo). Como estamos interessados em ter um indicador que seja atualizado em tempo real, o ideal seria que tivéssemos o histórico de revisões das séries para realizar nossa estimacão. Dessa forma, seria possível ver o quanto as séries escolhidas aumentam a previsibilidade do PIB no momento em que são divulgadas e não como elas se correlacionam *ex-post*, já que diversas séries são revisadas após a divulgação do PIB de forma a garantir sua coerência.

Infelizmente, o histórico de revisões está disponível apenas para um número bastante reduzido de séries. Assim, foi feita a opção de se utilizar a versão atual das séries históricas para se fazer a estimacão com o maior número possível de séries. Além disso, como ocorreram mudanças metodológicas na série do PIB ao longo desse período, seria razoável de se esperar que possa ter ocorrido alguma quebra nas relações que as variáveis mensais tinham com o PIB. Dessa maneira, é possível que, mesmo que se tivesse acesso

ao histórico de revisões, a estimação usando as versões atuais de cada série gerassem previsões mais precisas para o *nowcasting* de PIB.

Agora, alguns comentários sobre as datas de divulgação. Como o modelo é feito para ser atualizado a cada nova divulgação de um dado, o ideal seria ter a data de divulgação para cada observação de cada série. Essa é uma informação que é relativamente fácil de se obter para observações mais recentes. Porém, ela é muitas vezes inexistente para as observações mais antigas de séries mais longas. Dessa forma, ao invés de se coletar o dia específico em que cada observação foi divulgada, foi usada a semana do mês em que a série é frequentemente divulgada. Assim, ao invés de ser atualizado a cada nova divulgação dos dados, o modelo é atualizado semanalmente. A vantagem desse método é que a semana do mês em que o dado é divulgado costuma se manter constante ao longo do tempo. Dessa maneira, é possível ter uma estimativa razoável de quando o dado foi divulgado mesmo quando não se tem o dia exato. Por fim, como vale dizer que, como estamos usando as versões atuais das séries e não os dados com o histórico de revisões, é incerto o ganho que o uso de todas as datas de divulgação traria (já que a data de divulgação está não está relacionada ao valor atual que consta na série, mas sim a um valor passado que não está sendo usado na estimação).

É válido comentar que as duas simplificações adotadas na confecção da base de dados (usar a versão atual dos dados e não o histórico de revisão e usar a semana de divulgação e não o dia exato) são as mesmas utilizadas em outros trabalhos similares, como por exemplo Gianonne, Reichlin e Small (2008).

Os dados foram compilados utilizando o SGS do Banco Central do Brasil e a plataforma Datastream da Thomson Reuters. O Datastream disponibiliza as datas futuras em que novas observações estarão disponíveis. Com base nessa informação foi possível se ter uma estimativa da semana de divulgação de cada série. Para séries que não estão disponíveis no Datastream, a fonte primária do dado foi procurada no SGS e, a partir de uma pesquisa no website de cada uma dessas fontes, foi estabelecida uma estimativa para a semana de divulgação.

Na Tabela 1 abaixo, segue uma enumeração de todas as séries que estão incluídas no modelo no momento. Além disso, o quadro tem a classe de dados a que a série se enquadra (as séries foram divididas em 5 classes: preços, setor externo, setor financeiro, empregos e salários e outros) e a semana do mês em que novas observações são

divulgadas. É válido notar que a coletânea de estatísticas pode ser alterada à medida que novas metodologias vão sendo desenvolvidas e novas séries criadas a partir delas. A metodologia do modelo permite acomodar os novos dados, de forma que o modelo não se torne obsoleto ao longo do tempo.

Tabela 1

Série	Classe	Semana de Divulgação
Índice geral de preços-disponibilidade interna (IGP-DI) - Var. % mensal	Preços	1
Índice de preços ao produtor amplo - Var. % mensal	Preços	1
Importações: Total	Externo	1
Importações: Bens de Consumo Duráveis	Externo	1
Importações: Bens de Consumo Não-Duráveis	Externo	1
Importações: Produtos Intermediários e Materiais Brutos	Externo	1
Importações: Bens de Capital	Externo	1
Importações: Combustíveis	Externo	1
Exportações: Total	Externo	1
Exportações: Produtos Primários	Externo	1
Exportações: Produtos Manufaturados	Externo	1
Exportações: Produtos Semi-Manufaturados	Externo	1
Exportações: Operações Especiais	Externo	1
Taxa de Câmbio (Real-Dólar)	Financeiro	1
Utilização de Capacidade - Manufatura	Outros	1
Produção Industrial	Outros	1
Produção Industrial - Manufatura	Outros	1
Índice nacional de preços ao consumidor-amplo (IPCA) - Var. % mensal	Preços	2
Índice nacional de preços ao consumidor-Amplo (IPCA) - Comercializáveis - Var. % mensal	Preços	2
Índice de Preços ao Consumidor-Amplo (IPCA) - Não comercializáveis - Var. % mensal	Preços	2
Índice nacional de preços ao consumidor-Amplo (IPCA) - Preços monitorados - Total - Var. % mensal	Preços	2
Índice nacional de preços ao consumidor-Amplo (IPCA) - Núcleo médias aparadas com suavização - Var. % mensal	Preços	2
Índice de Preços ao Consumidor-Amplo (IPCA) - Bens não-duráveis - Var. % mensal	Preços	2
Índice de Preços ao Consumidor-Amplo (IPCA) - Bens semi-duráveis - Var. % mensal	Preços	2
Índice de Preços ao Consumidor-Amplo (IPCA) - Duráveis - Var. % mensal	Preços	2
Índice de Preços ao Consumidor-Amplo (IPCA) - Serviços - Var. % mensal	Preços	2
Índice nacional de preços ao consumidor - Amplo (IPCA) - Núcleo por exclusão - Sem monitorados e alimentos no domicílio - Var. % mensal	Preços	2
Índice nacional de preços ao consumidor (INPC) - Var. % mensal	Preços	2
Índice de Quantum de Importações	Externo	2
Índice de Quantum de Importações - Bens de Consumo Duráveis	Externo	2
Índice de Quantum de Importações - Bens de Consumo Não-Duráveis	Externo	2
Índice de Quantum de Importações - Bens Intermediários	Externo	2
Índice de Quantum de Importações - Bens de Capital	Externo	2
Índice de Quantum de Importações - Combustíveis	Externo	2
Índice de Quantum de Exportações	Externo	2
Índice de Quantum de Exportações - Produtos Primários	Externo	2
Índice de Quantum de Exportações - Bens de Consumo Duráveis	Externo	2
Índice de Quantum de Exportações - Bens de Consumo Não-Duráveis	Externo	2
Índice de Quantum de Exportações - Bens Intermediários	Externo	2
Índice de Quantum de Exportações - Bens de Capital	Externo	2
Índice de Quantum de Exportações - Combustíveis	Externo	2
Índice de Quantum de Exportações - Produtos Manufaturados	Externo	2
Índice de Quantum de Exportações - Produtos Semi-Manufaturados	Externo	2
Reservas Internacionais	Externo	2
Índice de Preço de Exportações	Externo	2
Índice de Preço de Importações	Externo	2
CDI (% por mês)	Financeiro	2
Novos Registros de Veículos	Outros	2
Índice do Emprego Formal - Índice	Emprego e salários	3
Índice geral de preços do mercado (IGP-M) - Var. % mensal	Preços	4
Índice nacional de custo da construção (INCC) - Var. % mensal	Preços	4
Custo Unitários do Trabalho	Emprego e salários	4
Saldo em Conta Corrente	Externo	4
Conta Corrente - Bens (Crédito)	Externo	4
Conta Corrente - Bens (Débito)	Externo	4
Conta Corrente - Renda Primária (Crédito)	Externo	4
Conta Corrente - Renda Primária (Débito)	Externo	4
Conta Corrente - Serviços (Crédito)	Externo	4
Conta Corrente - Serviços (Débito)	Externo	4
Conta Corrente - Renda Secundária (Crédito)	Externo	4
Conta Corrente - Renda Secundária (Débito)	Externo	4
Índice de confiança do consumidor - São Paulo	Outros	4
M0	Financeiro	4
M1	Financeiro	4
M2	Financeiro	4
M3	Financeiro	4
M4	Financeiro	4
Ibovespa	Financeiro	4

IV - METODOLOGIA

A metodologia empregada aqui segue de perto o que foi realizado por Giannone, Reichlin e Small (2008) para a economia americana. Para facilitar a compreensão, a notação utilizada pelos autores foi mantida praticamente inalterada.

Primeiramente, é importante definir o que gostaríamos de fazer. Queremos a projeção do PIB em um trimestre q usando as n séries mensais disponíveis no mês v . Ou seja, gostaríamos de ter:

$$Proj[Y_q | \Omega_v^n]$$

Onde Y_q é o PIB no trimestre q e Ω_v^n representa o vetor com n séries mensais disponíveis até o mês v . Porém, o vetor Ω_v^n é composto por blocos $\Omega_v^n = \Omega_v^{n_0}, \Omega_v^{n_1}, \Omega_v^{n_2}, \dots$; em que $\Omega_v^{n_0}$ é o vetor de séries mensais no mês v com dados disponíveis até o mês v , $\Omega_v^{n_1}$ é o vetor de séries mensais no mês v com dados disponíveis até o mês $v - 1$ e assim por diante.

Dado o alto número de séries envolvidas em Ω_v^n , a estimação direta do modelo com todas as séries iria performar muito mal na previsão do PIB (dada a alta incerteza que existiria nos parâmetros estimados). Dessa forma, para solucionar esse problema, se utiliza um modelo dinâmico de fatores. A ideia dessa classe de modelos é que alguns fatores comuns latentes são responsáveis pelo movimento conjunto das várias séries temporais. Dessa forma, é possível resumir a grande maioria da informação disponível nas séries nesses fatores, gerando um modelo de estimação muito mais preciso e com uma previsão mais eficiente da série em questão (nesse caso, o PIB).

Para realizar a estimação desse modelo, é preciso primeiro realizar algumas transformações nas séries históricas em questão. Como o PIB é uma série trimestral, é preciso que no final de cada trimestre as séries mensais reflitam os seus valores trimestrais. Além disso, para a estimação do modelo de fatores, é preciso que as séries em questão sejam estacionárias. Dessa forma, as séries não-estacionárias presentes na base de dados são transformadas para sua versão estacionária.

Assim, assume-se a seguinte estrutura para a i -ésima (com $i = 1, 2, \dots, n$) série mensal em um modelo com r fatores:

$$x_{i,t|v_j} = \mu_i + \lambda_{i1}f_{1,t} + \dots + \lambda_{ir}f_{r,t} + \xi_{i,t|v_j}$$

Onde $x_{i,t|v_j}$ representa a série i , com t observações disponíveis no mês v_j , $f_{r,t}$ representa o r -ésimo fator com t observações disponíveis e $\xi_{i,t|v_j}$ é o componente idiossincrático da série (ortogonal com os fatores).

Resumindo as equações para as séries mensais em formato matricial temos:

$$X_{t|v_j} = \mu + \Lambda F_t + \xi_{t|v_j}$$

Onde $X_{t|v_j} = (x_{1t|v_j}, \dots, x_{nt|v_j})'$, $\xi_{t|v_j} = (\xi_{1t|v_j}, \dots, \xi_{nt|v_j})'$, $F_t = (f_{1,t}, \dots, f_{r,t})'$ e Λ é uma matriz $n \times r$ onde se encontram os elementos λ_{ij} .

Podemos então definir nossa previsão para o PIB dados as informações disponíveis, o *nowcasting* como:

$$\hat{y}_{3k|v_j} = \alpha + \beta' \hat{F}_{3k|v_j}$$

Onde $\hat{F}_{3k|v_j} = E[F_{3k} | \Omega_{v_j}; \omega]$ (esperança da matriz dos fatores para o terceiro mês do trimestre k dado o conjunto de informações, i. e. as séries mensais, no mês v_j e o modelo ω).

A estimação dos fatores comuns é feita normalmente a partir da análise dos componentes principais das séries. Porém, é importante lembrar que a estrutura de dados montada não é completamente contemporânea. Dessa forma, como a análise de componentes principais requer que as séries sejam todas contemporâneas, apenas o seu uso iria descartar informações importantes na margem (em particular, iria descartar as observações mais recentes, que, *a priori*, seriam as observações que carregam as informações mais atualizadas sobre a economia). Para resolver esse problema se usa uma estimação em duas etapas, o que Stock e Watson (2011) chamam de um método de estimação de fatores de terceira geração.

Na primeira etapa, é feita a estimação dos parâmetros usando uma regressão de MQO nos componentes principais extraídos de um painel dos dados truncado na data da última divulgação da série mais atrasada. Na segunda etapa, usamos os estimadores encontrados para aplicar o filtro de Kalman e obter nossa estimação final para os fatores.

Para o uso do filtro de Kalman, é preciso especificar um pouco mais o modelo usado. Primeiramente, define-se a estrutura dinâmica dos fatores como um VAR:

$$F_t = AF_{t-1} + Bu_t$$

Onde u_t é um ruído branco. Além disso, o componente idiossincrático $\xi_{t|v_j}$ também é um ruído branco e é ortogonal a u_t . Assim, podemos estimar os fatores (e sua precisão) usando o filtro de Kalman nas seguintes equações:

$$\hat{F}_{t|v_j} = E \left[F_t \mid \Omega_{v_j}; \omega_{\hat{\theta}} \right]$$

$$\hat{V}_{v_j} = E[(F_t - \hat{F}_t)(F_t - \hat{F}_t)'; \omega_{\hat{\theta}}]$$

Onde $\omega_{\hat{\theta}}$ representa o modelo com os parâmetros estimados por MQO. Dessa maneira, para a estimação, o filtro de Kalman está implicitamente fazendo uma previsão para as séries com observações faltando. Assim, toda vez que o modelo é atualizado, essas previsões implícitas também são alteradas. Dessa forma, as mudanças que ocorrem no valor previsto para o PIB decorrem não apenas da presença da inclusão de uma nova variável, mas também nas mudanças que essa variável causa nas previsões para os valores futuros da série.

Assim como Giannone, Reichlin e Small (2008), o modelo utiliza dois fatores comuns para fazer a estimação. Outras especificações foram tentadas, mas não geraram resultados significativamente diferentes.

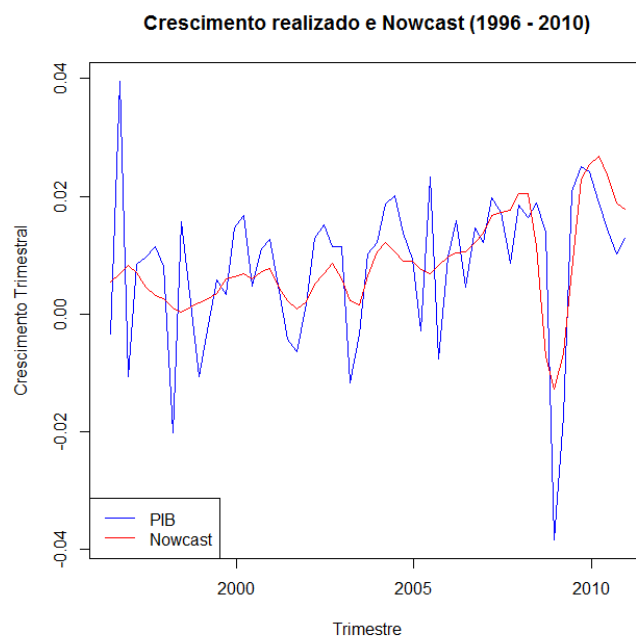
V - RESULTADOS

O modelo de *nowcasting* descrito na seção anterior foi estimado usando dados mensais de 1996 até o início de 2011 para as séries explicitadas no capítulo “Dados”. Então, o modelo foi estimado novamente ao final de cada semana para incorporar as informações novas a medida em que elas ficavam disponíveis.

Primeiramente, vamos analisar se o modelo consegue capturar bem o movimento do PIB durante sua janela de estimação. Porém, como explicado acima, o modelo foi estimado diversas vezes. Dessa forma, não existe um único modelo a ser analisado. Entretanto, analisar a sequência de todas as estimações também não traz muitas informações novas, já que diversas vezes as mudanças entre um modelo e outro são muito pequenas. Além disso, como a principal função do *nowcasting* é a previsão, a avaliação da precisão *in sample* dos modelos não é tão relevante. Portanto, a análise vai focar em dois pontos no tempo: a primeira estimação, com dados até o início de 2011 e a última estimação, com dados até o final de 2016.

O Gráfico 1 abaixo mostra os resultados para a primeira estimação. É importante notar que, como ainda não se tem o crescimento para o primeiro trimestre de 2011, os valores *in sample* vão até o final de 2010:

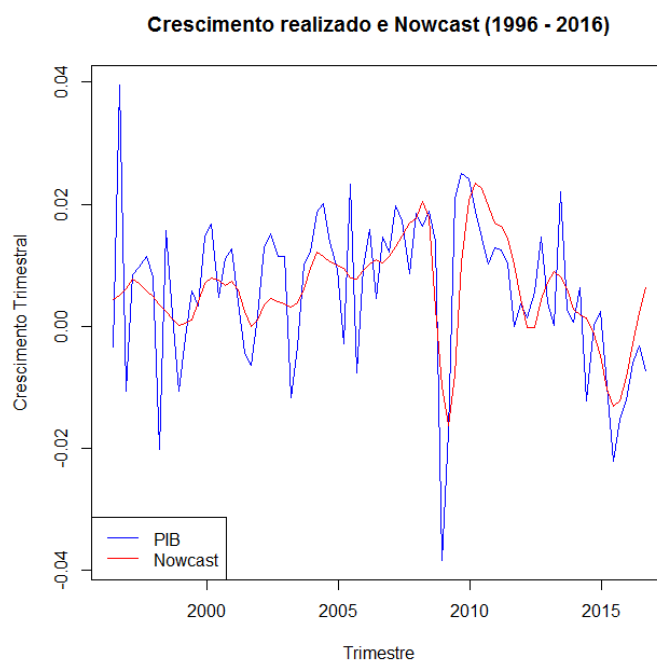
Gráfico 1



Podemos ver como o modelo captura bem a tendência para o crescimento trimestral, ainda que não reflita toda a variação presente trimestre a trimestre. Dessa forma, o R2 ajustado desse modelo é de 0,3466, relativamente baixo. Porém, como já foi comentado, a preocupação principal não é com a performance passada do modelo e sim se ele consegue gerar boas previsões para o trimestre seguinte, o que será analisado mais à frente.

Abaixo, o Gráfico 2 mostra os resultados para a última estimação do modelo, onde os valores *in sample* vão até o terceiro trimestre

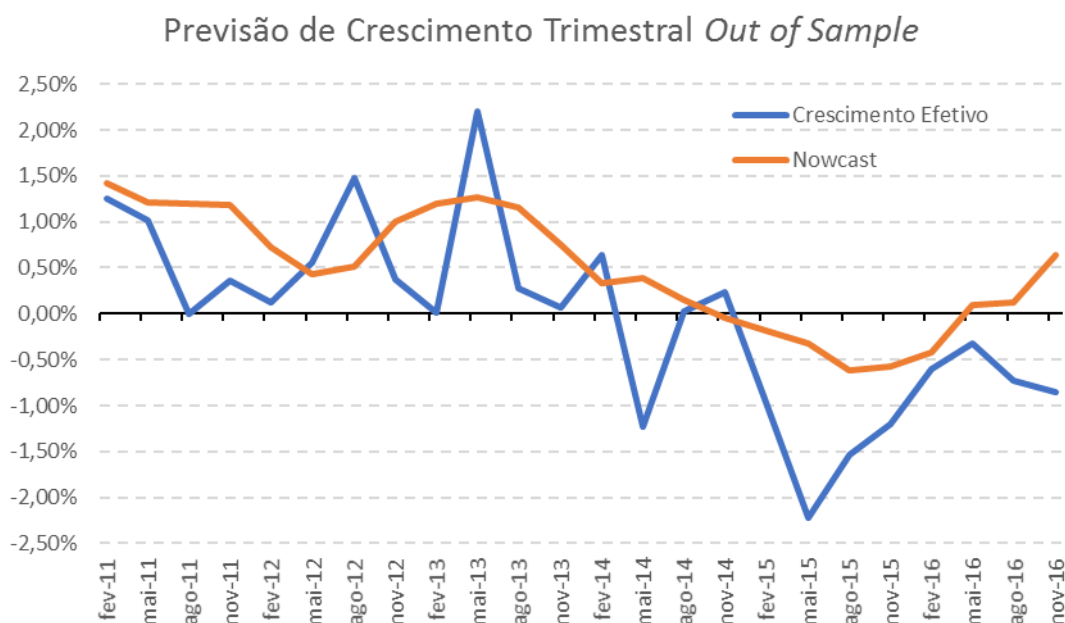
Gráfico 2



Os resultados são bastante similares ao que se tem quando analisamos a primeira estimação. Assim como naquele caso, o modelo consegue capturar bem a tendência do crescimento trimestral, ainda que não consiga capturar boa parte da variância do crescimento trimestral. Como na primeira estimação, o R2 ajustado é relativamente baixo (nesse caso, ele é de 0,3816).

Porém, é importante frisar novamente que a preocupação central do modelo é fazer a previsão para o crescimento no trimestre seguinte. Dessa forma, a principal preocupação é saber se o *nowcast* está realizando boas estimativas *out of sample*. O Gráfico 3, abaixo, mostra a evolução das estimativas *out of sample* do modelo contra o crescimento efetivo no mesmo período:

Gráfico 3



É importante entender o quão precisa são essas previsões *vis à vis* outras estimativas possíveis. Para isso, vamos comparar as previsões do modelo para o trimestre imediatamente seguinte com estimativas geradas de outras maneiras. Em particular, as outras previsões escolhidas são geradas a partir de: um modelo de crescimento constante, um passeio aleatório, um AR(1), a pesquisa FOCUS⁶ e um modelo de classe LASSO (usando as mesmas variáveis que o modelo de *nowcasting* utiliza e que são descritas no capítulo “Dados”).

Para realizar tal comparação, diversas métricas foram utilizadas. São elas: a média do erro absoluto (MAE), a raiz quadrada da média do erro quadrático (RMSE) e o erro máximo absoluto. Para todas as medidas, valores menores significam uma maior precisão das estimativas. A Tabela 2, abaixo, sumariza os resultados para o período de 2011 até 2016⁷:

⁶ É importante se fazer um esclarecimento em relação aos dados usados da pesquisa FOCUS. A série de PIB sofreu diversas revisões ao longo de sua história. Porém, as expectativas presentes no FOCUS são referentes à primeira divulgação do PIB. Porém, como explicado na seção “Dados”, está se usando a série atual de PIB para realizar o exercício. Portanto, é esperado que a performance do FOCUS reportada aqui seja diferente (provavelmente inferior) de sua performance real.

⁷ Para cada período, os parâmetros de cada um dos modelos eram estimados novamente de forma a incorporar toda a informação disponível até o momento

Tabela 2

Precisão Final Out Of Sample (2011-2016)						
Métrica	Crescimento Constante	Passeio Aleatório	AR(1)	Focus	LASSO	Nowcast
MAE	0,010	0,007	0,009	0,006	0,006	0,007
RMSE	0,012	0,010	0,011	0,007	0,008	0,009
Erro Máximo Absoluto	0,029	0,022	0,027	0,017	0,018	0,019

A partir da tabela, podemos observar como o modelo de *nowcast* gera estimativa mais precisas que modelos estatísticos ingênuos (crescimento constante, passeio aleatório e AR(1)). Porém, o modelo não consegue ser tão preciso quanto as estimativas baseadas no julgamento de especialistas (FOCUS) ou um outro modelo estatístico complexo (LASSO).

Uma característica importante do Brasil no período analisado, porém, é a forte recessão dos anos de 2015 e 2016. Recessões são muitas vezes difíceis de capturar estatisticamente. Além disso, períodos como o que o Brasil viveu podem ser marcados por choques estruturais e/ou mudanças de regime, de forma que um modelo estimado com dados passados pode não ser a melhor maneira de prever o crescimento para essa janela. Assim, também é válido comparar como os modelos estão se saindo em tempos “normais”, ou seja, no período de 2011 até 2014. Abaixo, a Tabela 3 compara as estimativas para esse período:

Tabela 3

Precisão Final Out Of Sample (2011-2014)						
Métrica	Crescimento Constante	Passeio Aleatório	AR(1)	Focus	LASSO	Nowcast
MAE	0.007	0.008	0.007	0.007	0.005	0.007
RMSE	0.008	0.010	0.008	0.008	0.007	0.008
Erro Máximo Absoluto	0.020	0.022	0.020	0.017	0.014	0.016

Como era esperado, as previsões são mais precisas, para praticamente todos os modelos, quando se retira a recessão recente da amostra de comparação. Além disso, a diferença que existia entre os modelos ingênuos e os modelos mais sofisticados praticamente desaparece, com a exceção do LASSO, que continua dando estimativas mais precisas que os modelos ingênuos.

As tabelas acima mostram os resultados para a última estimativa de cada um dos modelos. Porém, é importante frisar que essa não é a única métrica relevante. É importante também entender o quão rápido conseguimos chegar em uma previsão precisa.

Dessa forma, é importante comparar também as previsões iniciais de cada um dos modelos. As Tabelas 4 e 5, abaixo, contém tal comparação:

Tabela 4

Precisão Inicial Out Of Sample (2011 - 2016)						
Métricas	Crescimento Constante	Passeio Aleatório	AR(1)	Focus	LASSO	Nowcast
MAE	0,010	0,007	0,009	0,007	0,007	0,007
RMSE	0,012	0,010	0,011	0,009	0,008	0,009
Erro Máximo Absoluto	0,029	0,022	0,027	0,018	0,021	0,018

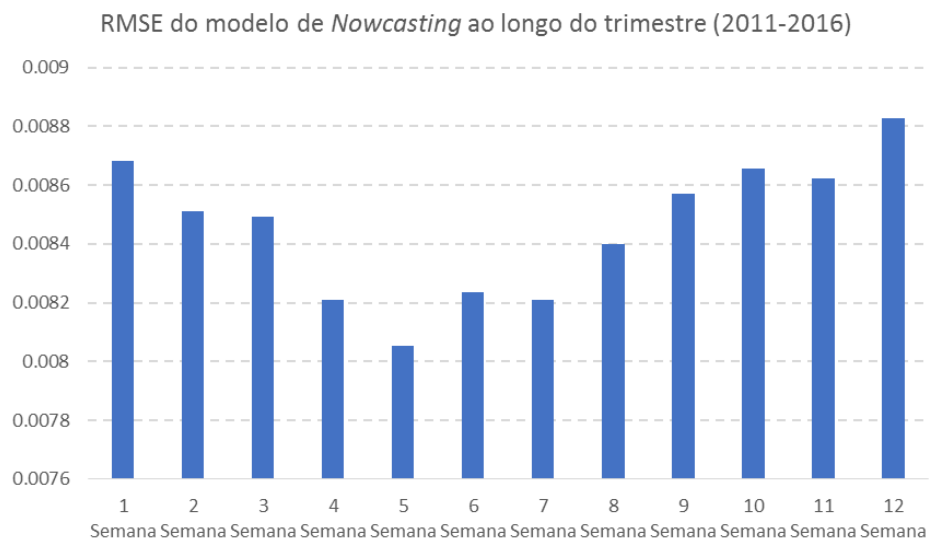
Tabela 5

Previsão Inicial Out Of Sample (2011 - 2014)						
Métricas	Crescimento Constante	Passeio Aleatório	AR(1)	Focus	LASSO	Nowcast
MAE	0.007	0.008	0.007	0.006	0.006	0.007
RMSE	0.008	0.010	0.008	0.008	0.007	0.008
Erro Máximo Absoluto	0.020	0.022	0.020	0.015	0.016	0.015

Primeiramente, como era o esperado, avaliando as estimativas do período entre 2011-2016, as previsões usando o LASSO e a pesquisa FOCUS, pioraram em relação a estimativa inicial (as previsões para crescimento constante, passeio aleatório e para o AR(1) são iguais nos dois períodos e por isso permanecem inalteradas). Porém, interessante, a previsão usando a técnica de *nowcasting* está tão precisa quanto sua estimativa final. Dessa forma, quando comparamos o conjunto de previsões iniciais, o modelo de *nowcasting* é essencialmente tão preciso quanto a pesquisa FOCUS e o modelo LASSO. Olhando para os resultados onde exclui-se a recessão, as conclusões são bastante similares e, assim como na avaliação das estimativas finais, o LASSO continua parecendo ligeiramente melhor que os outros modelos.

As tabelas acima apontam para uma outra questão importante a ser estudada, a precisão das estimativas do *nowcasting* ao longo do tempo. É importante lembrar que o modelo gera 12 estimativas entre a divulgação do crescimento de um trimestre e outro (são aproximadamente 12 semanas entre uma divulgação e outra). O esperado é que ao longo do trimestre a previsão fique mais precisa, ou seja, que o RMSE diminua ao longo do tempo. O Gráfico 4, abaixo, mostra a evolução do RMSE do modelo ao longo de cada uma das 12 semanas:

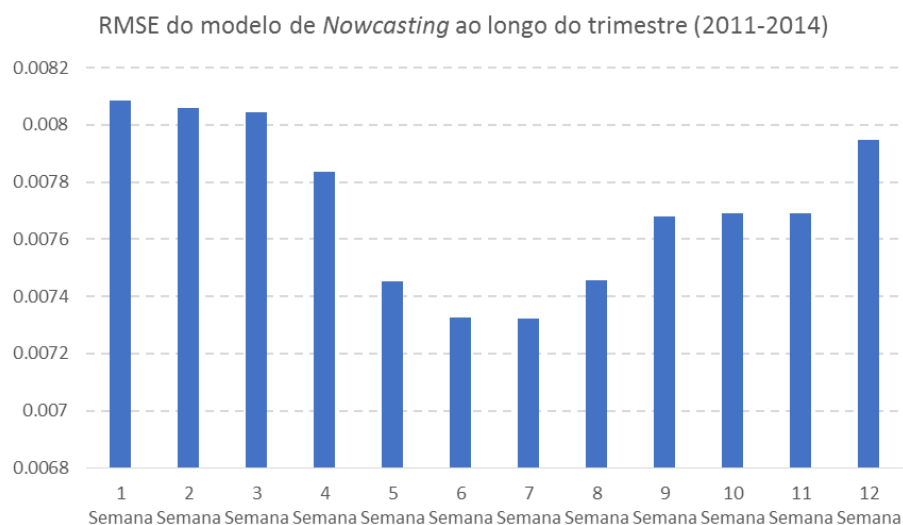
Gráfico 4



Diferentemente do que ocorre em Gianonne, Reichlin e Small (2008), o modelo de *nowcasting* para o Brasil não converge para sua melhor previsão ao longo do tempo. A precisão aumenta durante o primeiro mês de previsão para depois voltar a diminuir. Uma possível explicação para tal anomalia é a baixa quantidade de observações *out of sample*, não dando oportunidade para o RMSE convergir para seu valor real.

Vale ressaltar que essa anomalia não é causada pelo período de recessão entre 2015 e 2016. O Gráfico 5, abaixo, reproduz o RMSE ao longo do tempo para o período entre 2011 e 2014:

Gráfico 5



A tendência das estimativas ao longo do tempo permanece inalterada, isto é, as previsões melhoram ao longo do primeiro mês e voltam a piorar ao longo do terceiro. Porém, a melhor semana de previsão se altera (muda da 5ª semana para a 7ª semana) e o ritmo de melhora e piora também. Isso demonstra como algumas poucas observações podem fazer uma grande diferença para a trajetória apresentada por esse gráfico, o que pode sugerir que realmente o baixo número de observações *out of sample* é a razão pela qual a anomalia ocorre.

Isso, por fim, nos traz ao principal problema do estudo. A baixa quantidade de observações *out of sample* não polui apenas os resultados para a evolução das previsões do *nowcasting* ao longo do tempo, mas também a avaliação feita entre os diferentes modelos. Dessa forma, é perfeitamente possível que, caso existissem mais observações *out of sample* a classificação da precisão entre os modelos seria diferente. Porém, dado o tamanho da série completa de PIB para o Brasil e baixa quantidade de séries mensais disponíveis para períodos anteriores a 1996, análises que tivessem mais observações *out of sample* iriam esbarrar no problema de estimações imprecisas, dado que a janela de estimação seria pequena.

VI - CONCLUSÃO

Essa monografia procurou utilizar a técnica descrita em Giannone, Reichlin e Small (2008) para prever o crescimento trimestral do Brasil. Mostrou-se que o modelo tem uma performance preditiva boa quando comparada a outros tipos de modelo de previsão, ainda que não seja a melhor alternativa possível. Diferentemente do que ocorre no caso dos EUA, a precisão do modelo de *nowcast* não melhora de monotonicamente com o tempo. Porém, como a quantidade de observações *out of sample* é baixa, é possível que os resultados obtidos não reflitam perfeitamente a precisão do modelo *vis à vis* outras alternativas existentes.

Uma extensão possível é a possibilidade de adaptar o modelo algumas realidades do caso brasileiro. Em particular, uma técnica que permita utilizar séries que tenham início posterior ao de outras séries do painel permitiria incorporar várias estatísticas que podem ser possivelmente relevantes para a previsão do PIB no Brasil, como o IBC-Br, o desemprego e as vendas no varejo.

VII – REFERÊNCIAS BIBLIOGRÁFICAS

AASTVEIT, Knut Are; TROVIK, Torres G.; 2008. *Nowcasting Norwegian GDP: the role of asset prices in a small open economy*; Norges Bank Working Paper

ANGELINI, Elena; CAMBA-MENDEZ, Gonzalo; GIANNONE, Domenico; REICHLIN, Lucrezia; RÜNSTLER, Gerhard; 2011. *Short-term forecast of Euro area GDP growth*; ECB Working Paper

BANBURA, Marta; GIANNONE, Domenico; REICHLIN, Lucrezia; 2011. *Nowcasting*; Oxford Handbook of Economic Forecasting

BARHOUMI, Karim; BENK, Szilard; CRISTADORO, Riccardo; DEN REIJER, Ard; JAKAITIENE, Audrone; JELONEK, P.; RUA, António; RÜNSTLER, Gerhard; RUTH, K.; VAN NIEUWENHUYZE, C.; 2009. *Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise*; Journal of Forecasting 7, vol. 28, pág 595-611

DOZ, Catherine; GIANNONE, Domenico; REICHLIN, Lucrezia 2009 *A two-step estimator for large approximate dynamic factor models based on Kalman filtering*; Paris School of Economics Working Paper

FORONI, Claudia; MARCELLINO, Massimiliano; 2013. *A survey of econometric methods for mixed frequency data*; Norges Bank Working Paper

GIANNONE, Domenico; REICHLIN, Lucrezia; SMALL, David; 2008. *Nowcasting: The real-time informational content of macroeconomic data*; Journal of Monetary Economics 55, pág. 665-676

HIGGINS, Patrick; 2014. *GDPNow: A Model for GDP “Nowcasting”*; FRB Atlanta Working Paper

HIGGINS, Patrick; 2016. *GDPNow and Then*; macroblog;
<http://macroblog.typepad.com/macroblog/2016/05/gdpnow-and-then.html>

MATHESON, Troy; 2010. *An analysis of the informational content of New Zealand data releases: the importance of business opinion surveys*; RBNZ Working Paper

MODUGNO, Michelle; 2011. *Nowcasting inflation using high frequency data*; ECB Working Paper

STOCK, James; WATSON, Mark; 2011. *Dynamic Factor Models*; Oxford Handbook of Economic Forecasting

TUSELL, Fernando; 2011. *Kalman Filtering in R*; Journal of Statistical Software 2, vol 39

YIU, Matthew S.; CHOW, Kenneth K.; 2011. *Nowcasting Chinese GDP: information content of economic and financial data*; Working Paper of the Hong Kong Institute for Monetary Research