



PREVISÃO PARA O CRESCIMENTO ANUAL DO PIB DO ESTADO DE SÃO PAULO

DOI: 10.12957/synthesis.2015.30468

Elcion Caiado Rocha Lima^{*}, Luiz Almeida^{**},
Maria Helena Machado Rocha Lima^{***}

Resumo: Neste artigo apresentamos e estimamos um modelo de previsão para o PIB anual do estado de São Paulo. O modelo utiliza as informações contidas em um painel não balanceado, composto por diversas variáveis com periodicidade mensal e que estão relacionadas com o PIB do estado. Utilizamos a metodologia proposta por Giannone, Reichlin, Small (2008) que extrai informações dos dados através de fatores dinâmicos. O modelo deve permitir obter uma boa estimativa para o crescimento anual do PIB do estado, para 2014, 2015 e até o segundo trimestre de 2016.

Palavras-chave: São Paulo; Fatores Dinâmicos; Painel Balanceado; Componentes Principais.

Forecasts of GDP growth for the state of São Paulo

Abstract: This article estimates a model to forecast the annual growth of GDP of the State of São Paulo. The model uses data of an unbalanced panel which contains several variables with monthly frequency and related to the State's GDP. We use the methodology developed by Giannone, Reichlin and Small (2008) that extracts information from the data through the estimation of dynamic factors. The model may have a good performance when forecast the State of São Paulo GDP for 2014, 2015 and until the second quarter of 2016.

Keywords: São Paulo; Dynamic Factors; Balanced Panel; Major Componentes.

Classificação JEL: C13, C22, C33, C38 e C55.

* Professor Associado da UERJ.

** Aluno do Mestrado em Ciências Econômicas da UERJ.

*** Técnica aposentada do Cetem (Centro de Tecnologia Mineral) – Ministério da Ciência, Tecnologia e Inovação.

1 INTRODUÇÃO

As políticas governamentais são em grande parte baseadas em uma avaliação de qual é a evolução passada, corrente e futura do nível de atividade — em geral representado pelo PIB — em determinada região ou estado da federação. Infelizmente os dados do PIB anual do estado de São Paulo, parte do sistema de Contas Regionais do IBGE, só estão disponíveis com uma defasagem de dois anos.

Neste artigo apresentamos e estimamos um modelo de previsão para o PIB anual do estado de São Paulo. O modelo permite obter uma boa estimativa para o crescimento do PIB do estado, para os dois anos anteriores ao ano corrente (ano de referência) e até o terceiro trimestre do ano corrente.

Um conjunto de variáveis mensais relacionadas com componentes do PIB do estado de São Paulo, cujos valores estão disponíveis com diferentes defasagens no tempo, é utilizado na previsão do PIB anual do estado. Há ainda um número de anos para os quais existe informação sobre o PIB anual estadual. O modelo utiliza as informações contidas em um painel não balanceado, composto por diversas variáveis que apresentam periodicidade mensal, com informações disponíveis do início da amostra até diferentes momentos do tempo (até o mês corrente, até o mês passado etc.). Essas informações, contidas no painel não balanceado e os valores observados do PIB anual estadual, são utilizadas pelo modelo para prever o PIB anual para os dois anos anteriores ao ano corrente (2016) e para o ano corrente. Nas previsões para o PIB anual, de cada um dos dois anos anteriores ao ano corrente, são utilizadas informações mensais até dezembro do respectivo ano. Já a previsão para o ano corrente pode ser efetuada, em cada mês do ano corrente, utilizando toda a informação disponível do início da amostra até o respectivo mês. Neste artigo as informações vão até o mês de junho de 2016.

Dado o pequeno número de observações para o PIB anual, torna-se necessária a utilização de um modelo parcimonioso (com um número restrito de parâmetros a serem estimados). No modelo estimado a parcimônia é obtida estimando-se um pequeno número de fatores comuns que irão sumariar toda a informação contida no painel de dados não balanceados. Os modelos de fatores têm sido empregados com bastante sucesso nas previsões de variáveis macroeconômicas. Este sucesso vem sendo documentado por diversos autores.¹ O artigo está organizado em mais quatro seções além desta introdução: na seção 2 descrevemos a metodologia que permite a utilização de um painel não balanceado, na seção 3 descrevemos o modelo adotado e o método de estimação, na seção 4 descrevem-se os dados adotados e concluímos na seção 5.

2 A METODOLOGIA: UMA ABORDAGEM GERAL QUE PERMITE A UTILIZAÇÃO DE UM PAINEL NÃO BALANCEADO DE DADOS MENSAIS PARA PREVER O PIB ANUAL²

Através de um exemplo simples ilustraremos qual é a abordagem geral que permite utilizar um painel não-balanceado de dados mensais para se prever o PIB anual.

Seja Ω_v^n o conjunto das informações relevantes que contém todas as observações disponíveis até o mês v de todas as n variáveis relevantes. Então desejamos computar a seguinte projeção:

$$\text{Proj}[Y_q/\Omega_v^n], \quad (1)$$

onde: Y_q é o PIB estadual no ano q .

Admita-se agora que Ω_v^n seja composto por dois blocos $[\Omega_v^{n1}, \Omega_v^{n2}]$ e que os valores para o mês v das variáveis contidas em Ω_v^{n1} estejam disponíveis no mesmo mês v e que os valores para o mês v das variáveis contidas em Ω_v^{n2} só estejam disponíveis com um mês de defasagem (ou seja, só no mês $v+1$ serão conhecidos os valores das variáveis contidas em Ω_v^{n2}). Estamos, portanto, admitindo que o conjunto das informações contidas em Ω_v^n é um painel de dados não balanceados (com o formato de uma aresta dentada).

De uma maneira alternativa, mais sintética, o conjunto informação anterior pode ser representado pela seguinte notação:

$$\Omega_v = \{X_{itv}; i = 1, \dots, n; t = 1, \dots, T_{iv}\} \quad (2)$$

O conjunto de dados é composto por n variáveis, X_{itv} , onde o i ($i = 1, \dots, n$) identifica a série de tempo e t ($t = 1, \dots, T_{iv}$) indica, admitindo-se que se está no mês v , cada mês para o qual há observação para a variável i (t vai do início da amostra até a última observação disponível).

Como as informações para o PIB estadual são anuais e as informações do painel de dados não balanceados é mensal, é necessário que se introduza uma notação adicional. Há valores para o PIB acumulado nos últimos 12 meses apenas no último mês do ano para o qual há estimativas do PIB. Se k ($k=1,2,3,\dots$) indica os anos com informação sobre o PIB estadual anual, então $q = 12k$ indica os meses (dezembro de cada ano) para os quais há informação sobre o PIB acumulado nos últimos 12 meses.

Ao invés do índice do PIB anual estadual, o modelo irá projetar a taxa de crescimento do PIB anual. Seja Y_{12k} a taxa anual de crescimento do PIB estadual no mês $12k$. Então

$$\hat{Y}_{12k/v} = E(Y_{12k}/\Omega_v; M), v = 12k-h. \quad (3)$$

onde M indica o modelo utilizado para obter a esperança de Y_{12k} e h indica que a previsão foi feita h passos à frente.

A equação (3) acima é denominada “equação ponte” pois utiliza informação mensal para obter previsões para o PIB anual. A incerteza a respeito das previsões é obtida por:

$$VY_{12k} = E[(\hat{Y}_{12k/v} - Y_{12k})^2; M] \quad (4)$$

3 O MODELO E O MÉTODO DE ESTIMAÇÃO

A ideia fundamental subjacente ao modelo é a de explorar a colinearidade das séries de tempo mensais, que compõem o nosso painel com dados mensais, sumariando toda a informação disponível através da estimação de poucos fatores comuns. Ou seja, especificaremos um modelo com fatores comuns.

Seja X_{itv} uma observação disponível para o mês t , para a série de tempo i , quando se está no mês v . Admita-se ainda que a série de tempo foi transformada de forma a mensurar

valores (ou taxa de crescimento) acumulados nos últimos 12 meses. Vamos admitir a seguinte estrutura dos fatores para as séries de tempo mensais transformadas:

$$X_{it/v} = \mu_i + \lambda_{i1}f_{1,t} + \dots + \lambda_{ir}f_{r,t} + \xi_{i,t/v}, \quad i = 1, \dots, n \quad (5)$$

onde μ_i é uma constante; $\chi_{it} \equiv \lambda_{i1}f_{1,t} + \dots + \lambda_{ir}f_{r,t}$ e $\xi_{i,t/v}$ são dois processos estocásticos estacionários e ortogonais.

Admitiremos que os processos $\chi_{it/v}$ (os componentes comuns) são funções lineares de poucos $r \ll n$ fatores comuns não-observáveis (f_{it}) que capturam quase todos os comovimentos na economia, enquanto os processos $\xi_{i,t/v}$ (os componentes idiossincráticos) são afetados por choques específicos das variáveis.

O modelo acima pode ser reescrito adotando-se uma notação matricial:

$$X_{t/v} = \mu + \Lambda F_t + \xi_{t/v} = \mu + \chi_t + \xi_{t/v} \quad (6)$$

onde $X_{t/v} = (X_{1t/v}, \dots, X_{nt/v})'$, $\xi_{t/v} = (\xi_{1t/v}, \dots, \xi_{nt/v})'$, $F_t = (f_{1t}, \dots, f_{rt})'$ e Λ é uma matriz $n \times r$ dos pesos dos fatores com entrada genérica dada por λ_{ij} .

Sob a hipótese adicional de que a taxa de crescimento do PIB e os indicadores mensais têm uma distribuição normal conjunta, é possível obter previsões da taxa de crescimento do PIB através da seguinte função linear, equação ponte, dos fatores comuns esperados:

$$\hat{Y}_{12k/v} = \alpha + \beta' \hat{F}_{12k/v} \quad (7)$$

onde $\hat{F}_{12k/v} = E[F_{12k}/\Omega_v; M]$, para $v = 12k$.

O método de estimação proposto aqui é o do estimador em dois passos estudado por Doz, Giannone, e Reichlin (2006) e aplicado por Giannone, Reichlin, e Sala (2004) na identificação de choques macroeconômicos. A abordagem combina a técnica de componentes principais com o filtro de Kalman. O suavizador de Kalman é usado para computar, recursivamente, o valor esperado dos fatores comuns estimados.

Para aplicar a técnica do Filtro de Kalman para extrair os fatores comuns é necessária uma melhor explicitação da estrutura do modelo. Primeiramente, a dinâmica dos fatores comuns é parametrizada como um vetor autorregressivo:

$$F_t = AF_{t-1} + Bu_t; \quad u_t \sim WN(0, I_q) \quad (8)$$

onde: B é uma matriz $r \times q$ de posto q , A é uma matriz $r \times r$ com todas as raízes do $\det(I_r - Az)$ fora do círculo unitário e u_t é um processo ruído branco com os choques dos fatores comuns.

Em um modelo como este, um número de fatores comuns (r) maior do que o número de choques comuns (q) pretende capturar as relações antecedentes e defasadas entre as variáveis ao longo do ciclo de negócios.

Os componentes idiossincráticos são parametrizados sob a hipótese de que sejam ruídos brancos ortogonais, em cujo caso:

$$E(\xi_{t/v} \cdot (\xi_{t/v})') = \Psi_{t/v} = \text{diag}(\Psi_{1,t/v}, \dots, \Psi_{n,t/v}); \quad (9)$$

$$E(\xi_{t/v} \cdot (\xi_{t-s/v})') = 0, \quad s > 0 \text{ para todo } v. \quad (10)$$

Também se supõe que $\xi_{t/v}$ seja ortogonal ao choque comum u_t :

$$E(\xi_{t/v} \cdot u_{t-s/v}') = 0 \text{ para todo } s \text{ e } v. \quad (11)$$

Como o painel de dados mensais não é balanceado há a necessidade de lidar com observações faltantes. Neste problema a variância dos componentes idiossincráticos é parametrizada da seguinte forma:

$$\Psi_{it/v} = \begin{cases} \varphi_i, & \text{se } X_{it/v} \text{ é observável;} \\ \infty, & \text{se } X_{it/v} \text{ não é observável.} \end{cases} \quad (12)$$

Supondo que todos os erros sejam Gaussianos, as equações (6) – (12) caracterizam completamente o modelo. Maiores detalhes sobre a estimação do modelo encontram-se nas próximas seções.

4 Os DADOS COLETADOS, OS DADOS UTILIZADOS E A TRANSFORMAÇÃO DOS DADOS

Foram coletadas nove séries mensais de indicadores de atividade para o estado de São Paulo: índice de volume de vendas do comércio varejista (período: 2000-01 a 2016-06; fonte: IBGE); índices de produção da indústria geral (período: 2002-01 a 2016-06; fonte: IBGE); consumo de energia elétrica total (MWh) (período: 2004-01 a 2015-12; fonte: EPE); índice de atividade econômica regional - IBCR (período: 2003-01 a 2016-05; fonte: BCB); receita dos estados e municípios com ICMS (período: 1995-01 a 2016-05; fonte MF-Cotepe); índice do nível de emprego formal (período: 1996-01 a 2016-03; fonte: MTE); transferências da União para estados e municípios (deflacionado pelo IPCA) (período: 1995-01 a 2015-12; fonte: MF-STN); índice de inadimplência (período: 2004-01 a 2016-06; fonte: BCB) e índice do PIB mensal (2010-01 a 2016-04; fonte: SEADE).

Foram também coletadas as taxas de variação anuais do PIB do estado de São Paulo feitas pelo IBGE, encadeando os dados de 1995 a 2010 com os de 2010 a 2013. Estes valores encontram-se na Tabela 1, apresentada a seguir:

Tabela 1- Taxas de variação anuais do PIB de SP.

ANO	Volume ajustado (IBGE)
1996	1,27%
1997	3,38%
1998	-1,20%
1999	-1,34%
2000	4,26%
2001	0,38%
2002	1,17%
2003	-0,36%
2004	6,08%
2005	3,54%
2006	3,96%
2007	7,41%
2008	5,89%
2009	-0,78%
2010	7,94%
2011	3,69%
2012	1,49%
2013	2,89%

Fonte: IBGE

Foram feitas diversas simulações do modelo, utilizando-se diversos conjuntos de dados mensais, já que o emprego de todo o dado disponível não necessariamente implica estarmos utilizando um conjunto de dados mais informativo. Isto porque, para diversas séries, há um número diferente de observações faltantes que devem ser preenchidas antes da estimação do modelo. Optamos, no final, pela escolha das seguintes 7 séries: Índice de volume de vendas no varejo, Índice do nível de emprego formal, Índice de produção industrial, Receita dos estados e municípios com ICMS, Consumo de energia elétrica total (MWh), Índice do PIB mensal (SEADE) e Índice de atividade econômica regional (IBCR).

Cada variável mensal escolhida foi transformada para obter um cálculo, da sua taxa de variação acumulada nos últimos doze meses, similar ao da taxa anual de variação do PIB do estado. Esta transformação, para cada mês t , foi feita da seguinte forma: i) para cada mês t obteve-se a razão entre a soma das últimas 12 observações mensais (incluindo a observação do mês t) e a soma das 12 observações mensais imediatamente anteriores às doze (12) mais recentes; ii) esta razão menos 1 foi multiplicada pelo número cem (100).

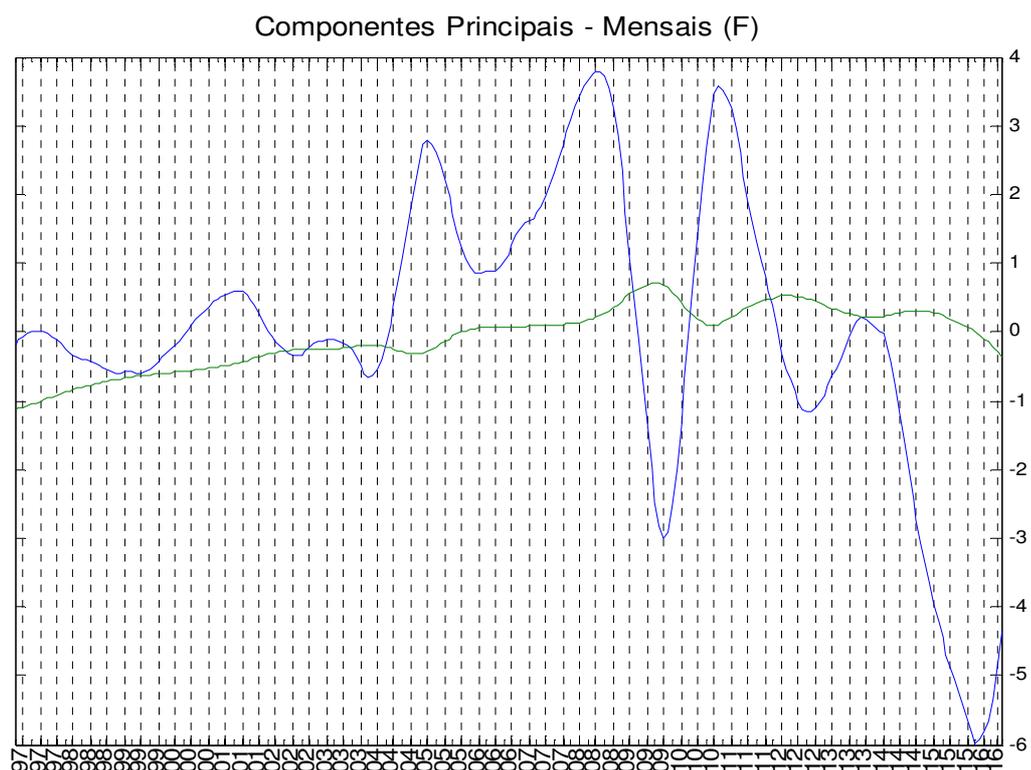
5 RESULTADOS

O modelo foi estimado sob a restrição de que o número de fatores comuns é igual a 2 ($r=2$) e de que o número de choques comuns é também igual a 2 ($q=2$). Veja a seção 3 para uma explicação detalhada do que é r e q . Foi utilizado o software MatLab para montar as rotinas que permitiram estimar o modelo para o estado de São Paulo.

O modelo foi estimado em dois passos. No primeiro, utilizando-se o filtro de Kalman, estimam-se os dois fatores comuns, com periodicidade mensal, dentro do período amostral que vai de janeiro de 1996 até junho de 2013. No segundo passo, utilizam-se os fatores comuns com periodicidade anual e a equação ponte (ver seção 2 e 3) para estimar as taxas de variação anuais do PIB.

A estimação dos fatores comuns no período amostral, juntamente com a estimação da equação ponte (equação 7), permite obter as taxas de variação anuais do PIB para os anos de 2014 e 2015. No Gráfico 1, apresentado a seguir, estão as estimativas dos fatores comuns mensais, utilizando o modelo da seção 3.

Gráfico 1



Nos Gráficos 2 e 3 são apresentados, respectivamente, os resíduos da equação-ponte e a comparação entre as taxas anuais de crescimento do PIB observadas e as estimadas pela equação ponte.

Gráfico 2

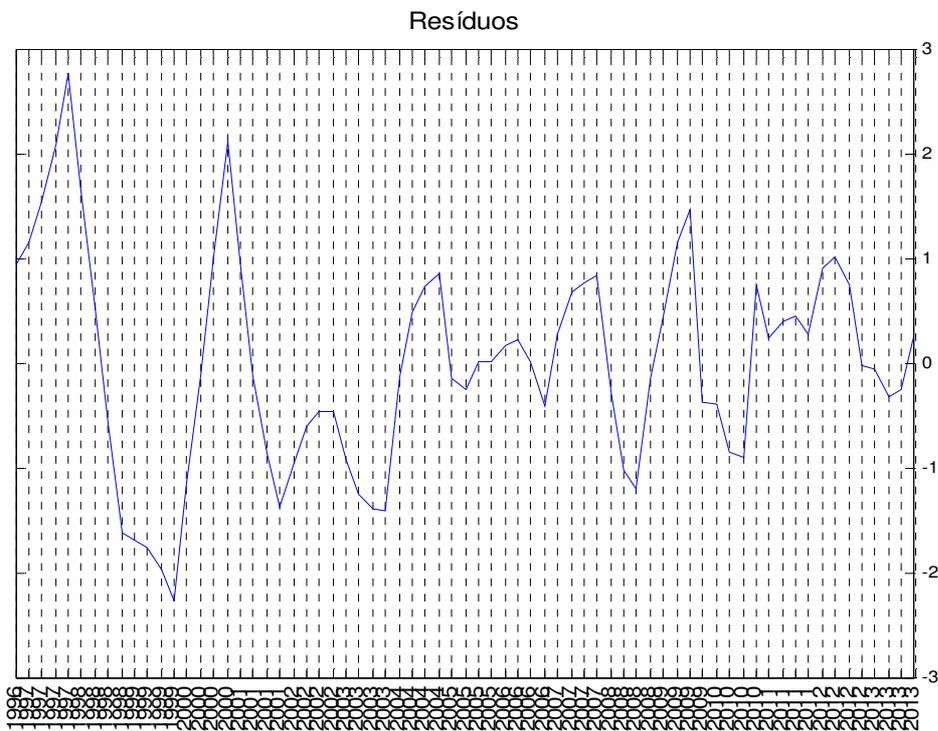
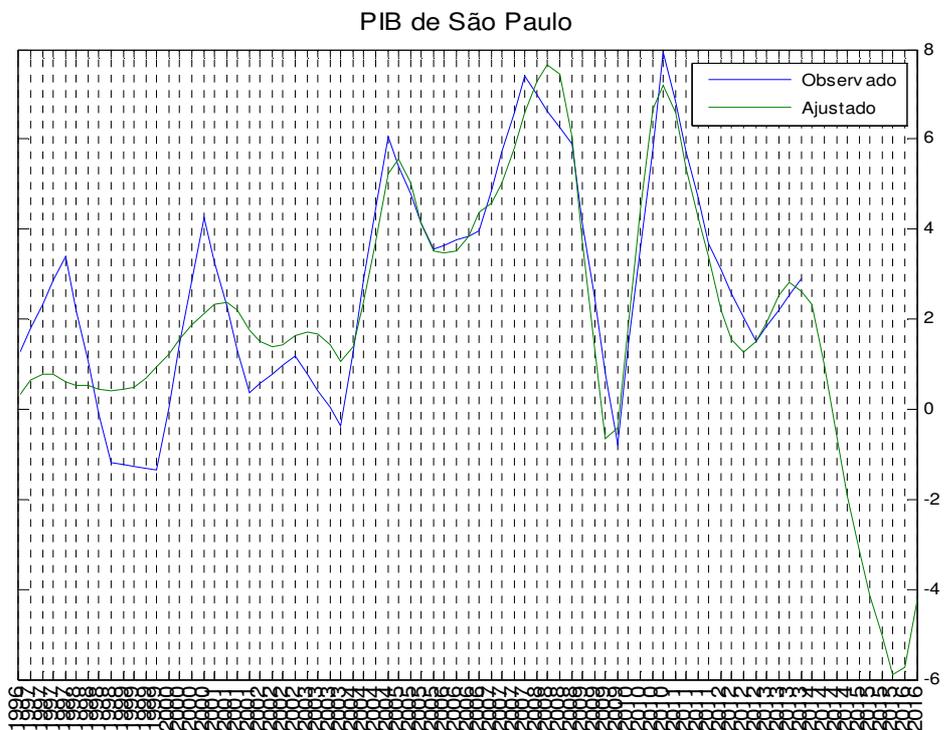


Gráfico 3



Na Tabela 2 são apresentadas as taxas de variação anual, do PIB do estado de São Paulo, observadas e ajustadas através da equação-ponte no período com informação sobre o crescimento do PIB (1996 a 2013).

Tabela 2

ANO	PIB IBGE	Nowcast SP
1996	1,27%	0,34%
1997	3,38%	0,63%
1998	-1,20%	0,42%
1999	-1,34%	0,94%
2000	4,26%	2,14%
2001	0,38%	1,77%
2002	1,17%	1,64%
2003	-0,36%	1,05%
2004	6,08%	5,22%
2005	3,54%	3,53%
2006	3,96%	4,37%
2007	7,41%	6,57%
2008	5,89%	6,04%
2009	-0,78%	-0,40%
2010	7,94%	7,20%
2011	3,69%	3,41%
2012	1,49%	1,52%
2013	2,89%	2,63%

Fonte: os autores.

Na Tabela 3, apresentada a seguir, estão as previsões para a taxa de crescimento anual do PIB do estado de São Paulo com frequência trimestral no período 2014:01 a 2016:03.

Tabela 3

Data	Taxa de variação anual (%)
1º Trimestre de 2014	2,321
2º Trimestre de 2014	0,973
3º Trimestre de 2014	-0,615
4º Trimestre de 2014	-1,915
1º Trimestre de 2015	-3,154
2º Trimestre de 2015	-4,131
3º Trimestre de 2015	-5,055
4º Trimestre de 2015	-5,914
1º Trimestre de 2016	-5,723
3º Trimestre de 2016	-4,209

Fonte: os autores.

NOTAS EXPLICATIVAS

¹ Como exemplos citem-se: Boivin e Ng (2005); Forni, Hallin, Lippi, e Reichlin (2005); D'Agostino e Giannone (2006); Giannone, Reichlin, e Sala (2004); Marcellino, Stock, e Watson (2003); Stock e Watson (2002a); Stock and Watson (2002b).

² Esta e a próxima seção seguem de perto a metodologia exposta no artigo de Giannone, D., L. Reichlin and D. Small (2008).

REFERÊNCIAS BIBLIOGRÁFICAS

- ALTISSIMO, F. et al. *EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle*. CEPR Discussion Papers 2001, DP 3108, Centre for Economic Policy Research, 2001.
- BAFFIGI, A.; GOLINELLI, R.; PARIGI, G. Bridge Models to Forecast the Euro Area GDP. *International Journal of Forecasting*, v.20, n.3, p. 447-460, 2004.
- BAI, J. Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, v.71, n.1, p.135-171, 2003.
- BAI, J.; NG, S. Determining the Number of Factors in Approximate Factor Models. *Econometrica*, v. 70, n.1, p. 191-221, 2002.
- BOIVIN, J.; NG, S. Understanding and Comparing Factor-Based Forecasts. *International Journal of Central Banking*, v. 1, n. 3, p. 117-151, 2005.
- CHOW, G. C.; LIN, A. Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series. *The Review of Economics and Statistics*, v. 53, n. 4, p. 372-75, 1971.
- CROUSHORE, D.; STARK, T. A Real-Time Data Set for Macroeconomists. *Journal of Econometrics*, v. 105, n. 1, p. 111-130, 2001.
- D'AGOSTINO, A.; GIANNONE, D. *Comparing Alternative Predictors Based on Large-Panel Dynamic Factor Models*. Working Paper Series 680, European Central Bank, 2006.
- D'AGOSTINO, A.; GIANNONE, D.; SURICO, P. *(Un) Predictability and Macroeconomic Stability*. Working Paper Series 605, European Central Bank, 2006.
- DOZ, C.; GIANNONE, D.; REICHLIN, L. *A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering*. Université Libre de Bruxelles, 2006. Unpublished manuscript.
- EVANS, M. D. Where Are We Now? Real-Time Estimates of the Macro Economy. NBER Working Paper 11064, *International Journal of Central Banking*, 2005.
- FED, C. *CFNAI Background Release*. Discussion paper. Disponível em: <<http://www.chicagofed.org/economicresearchanddata/national/pdffiles/CFNAIbga.pdf>>.
- FORNI, M. et al. Opening the Black Box: Structural Factor Models with large Cross-sections. Université Libre de Bruxelles, 2005. Manuscript.
- FORNI, M. et al. The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting. *Journal of the American Statistical Association*, v. 100, n. 471, p. 830-840, 2005.
- GIANNONE, D.; REICHLIN, L.; SALA, L. Monetary Policy in Real Time. In: GERTLER, M; ROGOFF, K. (Ed.). *NBER Macroeconomics Annual*. MIT Press, 2004. p. 161-200.
- GIANNONE, D.; REICHLIN, L.; SMALL, D. Nowcasting: The Real – Time Information Content of Macroeconomic Data. *Journal of Monetary Economics*, v. 55, n. 4, p. 665-676, 2008.
- KITCHEN, J.; MONACO, R. M. Real-Times Forecasting in Practice: The U.S. Treasury Staff's Real-Time GDP forecast System. *Business Economics*, p. 10-19, 2003.
- KOENIG, E. F.; DOLMAS, S.; PIGER, J. The Use and abuse of Real-Time Data in Economic Forecasting. *The Review of Economics and Statistics*, v. 85, n. 3, p. 618-628, 2003.
- MARCELLINO, M.; STOCK, J. H.; WATSON, M. W. Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information. *European Economic Review*, v. 47, n. 1, p. 1-18, 2003.
- ORPHANIDES, A. Monetary-Polity Rules and Great Inflation. *American Economic Review*, v. 92, n. 2, p. 115-120, 2002.
- RUNSTLER, G.; SÉDILLOT, F. *Short-Term Estimates Of Euro Area Real Gdp By Means Of Monthly Data*. Working Paper Series 276, European Central Bank, 2003.
- STOCK, J. H.; WATSON, M. W. Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, v. 97, n. 460, p. 147-162, 2002a.
- _____. Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics*, v. 20, n. 2, p. 147-162, 2002b.