

Mendonça, Mário Jorge; dos Santos, Cláudio Hamilton; Martins, Thiago Guerrera

**Working Paper**

## Aplicação de um modelo fatorial dinâmico para previsão da arrecadação tributária no Brasil

Texto para Discussão, No. 1453

**Provided in Cooperation with:**

Institute of Applied Economic Research (ipea), Brasília

*Suggested Citation:* Mendonça, Mário Jorge; dos Santos, Cláudio Hamilton; Martins, Thiago Guerrera (2009) : Aplicação de um modelo fatorial dinâmico para previsão da arrecadação tributária no Brasil, Texto para Discussão, No. 1453, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília

This Version is available at:

<https://hdl.handle.net/10419/91130>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# **TEXTO PARA DISCUSSÃO Nº 1453**

## **APLICAÇÃO DE UM MODELO FATORIAL DINÂMICO PARA PREVISÃO DA ARRECADAÇÃO TRIBUTÁRIA NO BRASIL**

**Mário Jorge Mendonça  
Cláudio Hamilton dos Santos  
Thiago Guerrera Martins**



# TEXTO PARA DISCUSSÃO Nº 1453

## APLICAÇÃO DE UM MODELO FATORIAL DINÂMICO PARA PREVISÃO DA ARRECADAÇÃO TRIBUTÁRIA NO BRASIL

**Mário Jorge Mendonça\***  
**Cláudio Hamilton dos Santos\*\***  
**Thiago Guerrera Martins\*\*\***

Rio de Janeiro, dezembro de 2009

---

\* Técnico de Planejamento e Pesquisa da Diretoria de Estudos e Políticas Macroeconômicas – DIMAC/Ipea.

\*\* Técnico de Planejamento e Pesquisa da DIMAC/Ipea.

\*\*\* Pesquisador-bolsista do Programa de Pesquisa para o Desenvolvimento Nacional (PNPD) na Coordenação de Finanças Públicas da DIMAC/Ipea.

## **Governo Federal**

**Secretaria de Assuntos Estratégicos da  
Presidência da República**  
**Ministro** Samuel Pinheiro Guimarães Neto

**ipea** Instituto de Pesquisa  
Econômica Aplicada

Fundação pública vinculada à Secretaria de Assuntos Estratégicos da Presidência da República, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

### **Presidente**

Marcio Pochmann

### **Diretor de Desenvolvimento Institucional**

Fernando Ferreira

### **Diretor de Estudos, Cooperação Técnica e Políticas Internacionais**

Mário Lisboa Theodoro

### **Diretor de Estudos e Políticas do Estado, das Instituições e da Democracia (em implantação)**

José Celso Pereira Cardoso Júnior

### **Diretor de Estudos e Políticas Macroeconômicas**

João Sicsú

### **Diretora de Estudos e Políticas Regionais, Urbanas e Ambientais**

Liana Maria da Frota Carleial

### **Diretor de Estudos e Políticas Setoriais, Inovação, Produção e Infraestrutura**

Márcio Wohlers de Almeida

### **Diretor de Estudos e Políticas Sociais**

Jorge Abrahão de Castro

### **Chefe de Gabinete**

Persio Marco Antonio Davison

### **Assessor-chefe de Comunicação**

Daniel Castro

## **TEXTO PARA DISCUSSÃO**

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

As opiniões emitidas nesta publicação são de exclusiva e de inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou da Secretaria de Assuntos Estratégicos da Presidência da República.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

URL: <http://www.ipea.gov.br>

Ouvidoria: <http://www.ipea.gov.br/ouvidoria>

ISSN 1415-4765

JEL: H20, H22, C32

# SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO	7
2 COMPOSIÇÃO DA CTBB E REVISÃO DA LITERATURA	8
3 MODELO FATORIAL DINÂMICO COM COMPONENTE SAZONAL	10
4 RESULTADOS DA APLICAÇÃO A COMPONENTES SELECIONADOS DA CTBB	14
5 COMENTÁRIOS FINAIS	15
REFERÊNCIAS	16
ANEXO A	18
ANEXO B	30



## **SINOPSE**

Este artigo tem por objetivo estimar um modelo fatorial dinâmico (MFD) bayesiano para análise e previsão de uma *proxy* da carga tributária no Brasil mensal no período 1996-2007. Argumenta-se que o emprego desse tipo de modelo é oportuno por permitir o tratamento conjunto do elevado número de tributos que compõem a carga tributária bruta brasileira (CTBB) – simultaneamente levando em consideração as informações contidas nas inter-relações existentes entre esses últimos e permitindo a identificação dos fatores subjacentes às dinâmicas dos mesmos. Além disso, e diferentemente do que é usual na literatura, o componente sazonal das séries é modelado endogenamente, permitindo a obtenção de estimativas melhor ajustadas aos dados e previsões mais confiáveis – uma vez que a sazonalidade é uma característica marcante das séries de arrecadação tributária. Por fim, um exercício de projeção para o ano de 2008 é realizado para os 20 impostos que compõem a nossa base de dados.

## **ABSTRACT**

The aim of this article is to estimate a Bayesian factorial dynamic model for the analysis and forecasting of the Brazilian tax burden (BTB) using monthly data from 1996 to 2007. Twenty taxes are responsible for about 80% of the BTB, each of which with a distinct seasonal pattern. The factorial model has no problems accommodating the high dimensionality of the data—contrarily to what happens, for instance, with VARs—while simultaneously allowing the identification of a short number of factors responsible for the joint dynamics of the various taxes. Therefore, this procedure allows one to obtain relevant insights about the public revenues in Brazil. Moreover, due to the fact that seasonality is a remarkable feature of the series of government receipts, the seasonal component is modeled endogenously using a Fourier form representation that is an unrestricted and flexible way to assess seasonality. Finally, we forecast the future path of the public receipts separately for the period of 2008.





# 1 INTRODUÇÃO

A carga tributária bruta brasileira (CTBB) é dada pela soma dos recursos que o governo recolhe compulsoriamente do setor privado. O aumento da CTBB eleva, geralmente,<sup>1</sup> os recursos à disposição do governo, permitindo a diminuição do endividamento público e/ou aumentos na oferta de bens e serviços públicos à população. A dinâmica da CTBB é, portanto, um determinante crucial das dinâmicas de diversas outras grandezas macroeconômicas relevantes.

Entretanto, não é trivial modelar o comportamento dinâmico da CTBB. Por um lado, ela é, na verdade, um (complexo) agregado de dezenas de tributos diferentes, cujas arrecadações dependem de diversas variáveis diferentes e, portanto, apresentam dinâmicas (e sazonalidades) sensivelmente diferentes. Além disso, a modelagem econométrica da CTBB (e de seus componentes) é muito dificultada pelas seguidas mudanças em alíquotas e bases de incidência de tributos preexistentes – e, vez por outra, também pela eliminação (temporária ou não) desses últimos e/ou a criação de novos tributos – que têm caracterizado a administração tributária do país no período pós-Plano Real.<sup>2</sup>

Embora pareça razoável – como primeira aproximação, pelo menos – supor que as arrecadações dos diferentes tributos que compõem a CTBB acompanhem o movimento geral da economia (ver MENDONÇA, PIRES e MEDRANO, 2008; SANTOS, RIBEIRO e GOBETTI, 2008, *inter alia*), é fácil perceber que as heterogeneidades existentes entre os mesmos fazem com que os resultados de modelos agregativos sejam usualmente pouco precisos. Por outro lado, as seguidas mudanças na legislação tributária<sup>3</sup> fazem com que haja bons motivos para se supor que as elasticidades e sazonalidades possam variar com o tempo. Assim, um instrumento econométrico que objetive modelar os vários componentes da CTBB deve não apenas lidar com a alta dimensionalidade inerente aos variados tipos de impostos, mas também saber tratar as características sazonais específicas a cada um deles.

Este estudo tem como objetivo a aplicação do modelo fatorial dinâmico (MFD) para previsão e análise da CTBB (usando dados mensais para o período 1995-2007) que tenta levar em consideração esses dois pontos cruciais. Uma razão fundamental para o uso do MFD é que este nos permite obter *insights* importantes por meio da redução de dimensionalidade (isto é, de aumento no grau de agregação) do modelo. Se, de algum modo, a informação contida nas inter-relações entre os vários tributos puder ser “passada para um conjunto menor de variáveis”, nossa habilidade de

---

1. Note-se que nem todo recurso que o governo recolhe compulsoriamente da sociedade é receita do governo. As contribuições para o Fundo de Garantia do Tempo de Serviço (FGTS), por exemplo, são recolhidas compulsoriamente da sociedade – e por isso entram no cálculo da CTBB – mas não são “receitas públicas” (visto que são de propriedade dos trabalhadores). Por outro lado, nem toda receita do governo é recolhida compulsoriamente da sociedade. Quando o cidadão paga entrada em um museu público, por exemplo, está aumentando a receita do governo, mas não a CTBB. Daí que aumentos na CTBB não necessariamente implicam aumentos nas receitas públicas e aumentos nessas últimas não necessariamente implicam aumentos na CTBB. Por outro lado, a maior parte das receitas públicas brasileiras é de natureza tributária e a maior parte da CTBB consiste em receitas públicas.

2. E, mesmo antes disso, como bem apontam Rezende, Oliveira e Araújo (2007).

3. Rezende, Oliveira e Araújo (2007, cap. 2) fazem um excelente resumo das inúmeras mudanças na legislação tributária brasileira no período 1998-2006.

identificar a “verdadeira” estrutura dos dados será melhor. Com efeito, quanto menor for a dimensionalidade do modelo, mais fácil será entendê-lo. Ademais, a incorporação das inter-relações entre as variáveis pelo modelo fatorial parece ter ajudado sensivelmente na qualidade do exercício de previsão que fizemos para o ano de 2008 (ver seção 4).

Notamos ainda que, diferentemente do que se costuma fazer na literatura (ver, por exemplo, SÁFADI e PENA, 2007), o componente sazonal das diferentes séries é modelado endogenamente, permitindo obter um melhor ajustamento e previsões mais precisas sobre a dinâmica futura dos impostos. Finalmente, também aplicamos procedimentos de inferência bayesiana (KIM e NELSON, 1999; WEST e HARRISON, 1997; CARTER e KOHN, 1994; GAMERMAN e LOPES, 2006; *inter alia*) que, entre outras vantagens, permitem resultados mais confiáveis em contextos em que a dimensão temporal da base de dados é pequena em relação ao espaço da dimensionalidade dos parâmetros (além de outras dificuldades, dificilmente tratáveis por meio de procedimentos “frequentistas”).

Até onde se sabe, o presente estudo é a primeira tentativa de aplicar MFDs aos dados tributários brasileiros. Talvez por conta das dificuldades mencionadas, poucos textos com especificações econométricas explícitas sobre esses dados foram publicados nos últimos 15 anos.<sup>4</sup> Esses textos diferem consideravelmente entre si – trabalhando com frequências, amostras e técnicas econométricas diferentes – mas nenhum deles procura explorar as informações contidas nas inter-relações entre as arrecadações dos vários tributos.

O restante deste artigo está estruturado da seguinte forma. A seção 2 apresenta alguns fatos estilizados acerca da composição da CTBB no período 1995-2008. Em seguida, a seção 3 apresenta uma descrição do modelo fatorial bayesiano que incorpora o componente sazonal de modo endógeno. Os resultados e a análise dos resultados do modelo são mostrados na seção 4. A seção 5 apresenta as principais conclusões do artigo.

## 2 COMPOSIÇÃO DA CTBB E REVISÃO DA LITERATURA

Conforme já assinalamos, a CTBB é composta de um número elevado de impostos. A seguir, faremos um resumo da estrutura tributária nacional.<sup>5</sup> Em termos gerais, a CTBB é formada por quatro grandes agregados de tributos, quais sejam: *i*) impostos sobre produtos; *ii*) outros impostos ligados à produção; *iii*) impostos sobre renda, patrimônio e capital; e *iv*) contribuições previdenciárias. São sete os principais “impostos sobre produto”: *i*) Imposto sobre Operações Relativas à Circulação de Mercadorias e sobre Prestações de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação (ICMS) (estadual); *ii*) Contribuição para Financiamento da Seguridade Social (Cofins); *iii*) Imposto sobre Produtos Industrializados (IPI); *iv*) Imposto sobre Serviços (ISS) (municipal); *v*) Imposto sobre Importações (II); *vi*) Imposto sobre Operações Financeiras (IOF); e *vii*) Contribuição

---

4. Hernández (1998), Portugal e Portugal (2001), Melo (2001), Carvalho (2001) e Siqueira (2002) deram todos importantes contribuições a essa literatura. Uma resenha desta literatura aparece em Santos, Ribeiro e Gobetti (2008).

5. Para uma descrição sucinta da estrutura tributária no Brasil, ver Santos, Ribeiro e Gobetti (2008).

de Intervenção no Domínio Econômico (Cide).<sup>6</sup> Apenas o ICMS e a Cofins, somados, responderam por mais de três quartos de todos os “impostos sobre produtos” e por pouco menos de um terço da CTBB total em 2006. Juntos os impostos sobre produtos – cujo valor está implícito no preço dos produtos e, portanto, muitas vezes passa despercebido pelo consumidor – respondem por pouco mais de 40% da CTBB.

São quatro os principais tributos brasileiros sobre a renda, o patrimônio e o capital – Impostos sobre o Rendimento das Pessoas Coletivas (IRPCs): o Imposto de Renda (IR); a Contribuição Social sobre o Lucro Líquido (CSLL) das pessoas jurídicas; o Imposto sobre a Propriedade de Veículos Automotores (IPVA) (estadual); e o Imposto Predial e Territorial Urbano (IPTU) (municipal). Apenas a arrecadação do IR – que é dividida em seis componentes, quais sejam: o IR pessoa física (IRPF); o IR pessoa jurídica (IRPJ); o IR retido na fonte (IRRF) sobre os rendimentos do trabalho; o IRRF sobre ganhos de capital; o IRRF sobre remessas de divisas ao exterior; e os outros tipos de IRRF – atingiu pouco menos de 60,5% da arrecadação total dos IRPCs, enquanto a CSLL, o IPTU e o IPVA representaram, respectivamente, 11,9%, 5,9% e 5,5% do total dos IRPCs nesse mesmo ano. Juntos os IRPCs representam pouco menos de 30% da CTBB.

As chamadas contribuições previdenciárias (CPs) respondem por cerca de um quarto da CTBB. Tal como ocorre com os demais componentes da CTBB, a lista de tributos relacionados às CPs é bastante extensa.<sup>7</sup> *Grosso modo*, porém, elas podem ser divididas em: *i*) contribuições de patrões e empregados para o Instituto Nacional do Seguro Social (INSS); *ii*) contribuições para o Programa de Integração Social (PIS)/ Programa de Formação do Patrimônio do Servidor Público (PASEP); *iii*) contribuições para o FGTS; e *iv*) contribuições para a aposentadoria dos funcionários públicos (pagas pelos próprios e pelo governo). A grande parte das mesmas incide sobre a folha de pagamentos de empresas e administrações públicas (isto é, sobre o “rendimento do fator de produção trabalho”).

Por fim, o último e quantitativamente menos importante componente da CTBB é o item “outros impostos ligados à produção”, composto basicamente por taxas diversas (por exemplo, “taxa de fiscalização das telecomunicações” ou “taxa de vigilância sanitária”) e por contribuições incidentes sobre a folha de pagamento de empresas e administrações públicas (como para o salário-família e para o Sistema “S”<sup>8</sup>).

Percebe-se, assim, a flagrante heterogeneidade dos componentes individuais da CTBB. A tabela 1 apresenta a amostra dos tributos usada neste estudo – composta por

---

6. Incidente sobre a importação e a comercialização de gasolina e seus derivados, *diesel* e seus derivados, querosene de aviação e outros querosenes, óleos combustíveis (*fuel-oil*), gás liquefeito de petróleo (GLP), inclusive o derivado de gás natural e de nafta, e álcool etílico combustível.

7. Ver IBGE (2004, anexo 8) para uma lista que exclui as contribuições para o PIS/PASEP (as quais foram reclassificadas como contribuições previdenciárias no Sistema de Contas Nacionais 2005).

8. Serviço Nacional de Aprendizagem Comercial (SENAC), Serviço Social do Comércio (SESC), Serviço Nacional de Aprendizagem do Cooperativismo (SESCOOP), Serviço Nacional de Aprendizagem Industrial (Senai), Serviço Social da Indústria (Sesi), Serviço Social do Transporte (SEST), Serviço Nacional de Aprendizagem do Transporte (SENAT), Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (Sebrae), Serviço Nacional de Aprendizagem Rural (Senar).

vários dos principais componentes dos quatro grandes tipos de tributo mencionados – bem como o percentual que cada um deles representa em relação ao total. Essa amostra se compõe basicamente de tributos federais, cuja base de dados é gerenciada pela Secretaria da Receita Federal (SRF), mas inclui também o ICMS estadual – com dados mensais disponibilizados pelo Conselho Nacional de Política Fazendária (Confaz). Estes 20 impostos representam pouco menos de 80% do total de tributos arrecadado pelas administrações públicas brasileiras.

TABELA 1  
**Descrição dos tributos**  
 (Em %)

Imposto	%	Imposto	%
ICMS	26,52	II	2,10
Previdência social	20,17	IRPF	1,30
Cofins	12,77	IPI – vinculado à importação	1,27
IRPJ – demais empresas	6,08	IOF	1,19
IRRF – rendimentos do trabalho	5,91	IRPJ – entidades financeiras	1,15
“Outros tributos”	4,64	IRRF – remessa ao exterior	1,09
CLPJ	3,68	IRRF – outros rendimentos	0,74
IRRF – rendimentos de capital	3,66	IPI – fumo	0,64
PIS/PASEP	3,58	IPI – automóveis	0,60
IPI – outros	2,36	IPI – bebidas	0,56

Fonte: Elaboração dos autores.

### 3 MODELO FATORIAL DINÂMICO COM COMPONENTE SAZONAL

O MFD pertence à classe de modelos de variáveis latentes, que são ferramentas importantes para a análise de dados multivariados. Uma razão dessa importância é a redução de dimensionalidade que esses modelos proporcionam. Se, de algum modo, a informação contida nas inter-relações de muitas variáveis puder ser passada, com boa aproximação, para um conjunto muito menor, nossa habilidade de ver a estrutura nos dados será melhor. Além disso, essa redução de dimensionalidade faz com que os modelos fatoriais sejam parcimoniosos, quando comparados com outros modelos que analisam dados multivariados, contudo não proporcionam essa redução de dimensionalidade, como os modelos VAR (HAMILTON, 1993), muito usados na literatura econométrica, mas que comportam apenas um conjunto restrito de variáveis.<sup>9</sup>

#### 3.1 DEFINIÇÃO DO MODELO

Suponha  $m$  variáveis relacionadas que obedecem à distribuição normal multivariada denotada por  $N(0, \Omega)$  onde  $\Omega$  denota a matriz  $m \times m$  não singular de variância-

9. Mais detalhes a respeito das vantagens do uso de modelos fatoriais podem ser encontrados em Bartholomew (1987).

covariância. Uma amostra aleatória de tamanho  $T$  é denotada por  $\{y_t, t = 1, \dots, T\}$ . O modelo utilizado nesse estudo é dado por

$$\begin{aligned} y_t &= \beta f_t + \underline{S}_t + e_t, \quad e_t \sim N(0, \Sigma) \\ \underline{f}_t &= \Gamma_1 \underline{f}_{t-1} + \Gamma_2 \underline{f}_{t-2} + \dots + \Gamma_p \underline{f}_{t-p} + \underline{w}_t, \quad \underline{w}_t \sim N(0, \Lambda) \end{aligned} \quad (1)$$

onde  $y_t$ ,  $\underline{f}_t$ ,  $e_t$  e  $\underline{w}_t$  são vetores de dimensão  $m$ ,  $k$ ,  $m$  e  $k$ , respectivamente;  $\beta$  é uma matriz  $m \times k$ ;  $\Gamma_i = \text{diag}(\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{ik})$ ;  $i = 1, \dots, p$ ;  $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2)$ ; e  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ .

Para algum inteiro positivo específico  $k \leq m$ , o modelo (1) relaciona cada  $y_t$  com um vetor  $k$ -dimensional de variáveis aleatórias  $\underline{f}_t$ , os fatores comuns, que obedecem a uma estrutura  $AR(p)$ , e um componente sazonal,  $\underline{S}_t = (S_{1t} \ S_{2t} \ \dots \ S_{mt})^T$  (WEST e HARRISON, 1997), definida por

$$S_{it} = \sum_{j=1}^6 \left[ a_{ij} \text{sen} \left( \frac{2\pi jt}{12} \right) + b_{ij} \cos \left( \frac{2\pi jt}{12} \right) \right] = \varphi_i X_t$$

onde:

$$X_t = \left( \text{sen} \left( \frac{2\pi t}{12} \right) \text{sen} \left( \frac{2\pi 2t}{12} \right) \dots \text{sen} \left( \frac{2\pi 6t}{12} \right) \cos \left( \frac{2\pi t}{12} \right) \cos \left( \frac{2\pi 2t}{12} \right) \dots \cos \left( \frac{2\pi 6t}{12} \right) \right)^T$$

e:

$$\varphi_i = (a_{i1} \ a_{i2} \ \dots \ a_{i6} \ b_{i1} \ b_{i2} \ \dots \ b_{i6}), \text{ para } i = 1, \dots, m$$

A ideia que motiva a aplicação da forma de Fourier para tratar sazonalidade está ligada ao fato de que qualquer padrão cíclico pode ser representado em termos da combinação linear de funções periódicas.<sup>10</sup> A principal motivação para o uso desta abordagem, além da flexibilidade, é a facilidade de se obter interpretação, já que as trajetórias em forma de ondas que se obtêm podem ser associadas aos variados padrões observados na prática.

Em problemas práticos, especialmente para valores grandes de  $m$ , o número de fatores  $k$  será pequeno quando comparado com  $m$ , de tal modo que muito da estrutura de variância-covariância é explicada pelos fatores comuns. As variâncias idiossincráticas,  $\sigma_i^2$ , medem a variabilidade residual em cada uma das variáveis dos dados, uma vez que a contribuição dos fatores já foi considerada.

10. Uma função  $g(t)$  é dita periódica se para algum inteiro  $p > 1$ , e todo  $n, t > 0$   $g(t + np) = g(t)$ .

### 3.2 RESTRIÇÕES NO MODELO FATORIAL

Um modelo com  $k$  fatores deve apresentar alguma restrição para ser definido como um modelo único, não apresentando problemas de identificação. Há várias maneiras de identificar o modelo adicionando restrições no  $\beta$ . Vamos seguir aqui a alternativa usada em Geweke e Zhou (1996), Aguilar e West (2000), entre outros, ao restringir  $\beta$  de modo que seja matriz bloco triangular inferior, assumida ser de posto completo, com elementos da diagonal principal iguais a um, como abaixo:

$$\beta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta_{2,1} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ \beta_{k,1} & \beta_{k,2} & \beta_{k,3} & \cdots & 1 \\ \beta_{k+1,1} & \beta_{k+1,2} & \beta_{k+1,3} & \cdots & \beta_{k+1,k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \beta_{m,1} & \beta_{m,2} & \beta_{m,3} & \cdots & \beta_{m,k} \end{pmatrix}$$

A restrição garante que  $\beta$  tenha posto completo  $k$ , resolvendo os problemas de identificação que surgem pela invariância do modelo sob mudanças de locação da matriz  $\beta$ . Ao mesmo tempo em que garante que o número de parâmetros livres no tempo  $t$  na representação fatorial não exceda os  $m(m+1)/2$  parâmetros de  $\Omega$  não restrito. Além de resolver os problemas de identificação, essa estrutura de  $\beta$  influencia a interpretação dos fatores e pode impactar no ajuste do modelo e na escolha do  $k$  em particular, mas não tem impacto nas previsões, pois as variâncias e covariâncias entre as séries são independentes da decisão de modelagem.

### 3.3 ESPECIFICAÇÃO DAS PRIORIS

Para completar a especificação do nosso modelo, temos que atribuir classes de prioris para os parâmetros  $f_0$ ,  $\beta$ ,  $\Gamma_i$ ,  $i = 1, \dots, p$ ,  $\Sigma$ ,  $\Lambda$  e  $\varphi_i$ ,  $i = 1, \dots, m$ . Seguem abaixo as prioris utilizadas:

- 1) Vetor de fatores comuns inicial
  - $f_0 \sim N(m_0, C_0)$
- 2) Elementos não nulos da matriz de cargas  $\beta$ 
  - $\beta_{ij} \sim N(\mu_0, V_0)$  para  $i > j$
- 3) Elementos da matriz diagonal de variância-covariância  $\Sigma$ 
  - $\sigma_i^2 \sim IG\left(\frac{n_\sigma}{2}, \frac{n_\sigma S_\sigma}{2}\right)$ , para  $i = 1, \dots, m$

4) Elementos das matrizes diagonais dos coeficientes autorregressivos  $\Gamma_1, \Gamma_2, \dots, \Gamma_p$

- $\gamma_{ij} \sim N(m_r, S_r)$ , para  $i = 1, \dots, p, j = 1, \dots, k$

5) Elementos da matriz diagonal de variância-covariância  $\Lambda$

- $\lambda_j \sim IG\left(\frac{n_\lambda}{2}, \frac{n_\lambda S_\lambda}{2}\right)$ ,  $j = 1, \dots, k$

6) Elementos dos vetores de coeficientes sazonais  $\varphi_1, \varphi_2, \dots, \varphi_m$ .

- $a_{ij} \sim N(m_s, C_s)$  e  $b_{ij} \sim N(m_s, C_s)$ , para  $i = 1, \dots, m$  e  $j = 1, \dots, 6$

### 3.4 MÉTODO MARKOV CHAIN MONTE CARLO (MCMC) PARA O MODELO FATORIAL COM $k$ FATORES

Nosso objetivo é obter as posteriores dos parâmetros do modelo (1), assim como fazer inferência sobre o processo fatorial  $\{f_t : t = 1, 2, \dots, T\}$ . Como uma solução exata requer o cálculo de integrais intratáveis do ponto de vista analítico, vamos usar o método Amostrador de Gibbs para obter uma aproximação das posteriores de interesse. As posteriores completas necessárias para implementar o algoritmo MCMC se encontram no anexo B.<sup>11</sup>

### 3.5 PREVISÃO DO MODELO FATORIAL DINÂMICO COM COMPONENTE SAZONAL

Vamos descrever nesta seção a metodologia para previsão  $h$  passos à frente, usando o modelo (1) com  $p = 1$  (que será o valor adotado em nossas análises, ver seção 4), a extensão para  $p > 1$  ocorre de forma simples e se encontra no anexo B.

Nosso interesse é na densidade preditiva  $h$  passos à frente:

$$p(y_{t+h} | y) = \int p(y_{t+h} | \tilde{f}_{t+h}, \beta, \varphi, \Theta) p(\tilde{f}_{t+h} | f_t, \beta, \varphi, \Theta) p(f_t | y) df_{t+h} df_t d\beta d\varphi d\Theta$$

onde:

$$\Theta = (\underline{\sigma}, \underline{\lambda}, \underline{\gamma}), \quad \underline{\sigma} = (\sigma_1, \dots, \sigma_m), \quad \underline{\gamma} = (\gamma_{11}, \dots, \gamma_{1k}), \quad \underline{\lambda} = (\lambda_1, \dots, \lambda_k)$$

$$(y_{t+h} | \tilde{f}_{t+h}, \beta, \varphi, \Theta) \sim N(\beta \tilde{f}_{t+h} + \varphi X_{t+h}, \Sigma) \text{ e } (f_t | f_t, \beta, \underline{\sigma}) \sim N(\underline{\mu}_h, V_h)$$

$$\text{com } \underline{\mu}_h = \Gamma_1^h f_t \text{ e } V_h = \sum_{k=1}^h \Gamma_1^{k-1} \Lambda (\Gamma_1^{k-1})^T, \text{ para } h > 0$$

11. Para maiores informações sobre como calcular tais condicionais completas no contexto de modelos fatoriais, sugere-se Kim e Nelson (1999).



Portanto, se  $\left\{ \left( \beta^{(1)}, f_{\tilde{t}}^{(1)}, \varphi^{(1)}, \Theta^{(1)} \right), \dots, \left( \beta^{(M)}, f_{\tilde{t}}^{(M)}, \varphi^{(M)}, \Theta^{(M)} \right) \right\}$  é uma amostra de  $p\left(f_{\tilde{t}}, \beta, \varphi, \Theta \mid y\right)$ , basta gerar uma amostra de tamanho  $M$  de  $\left(f_{\tilde{t+b}} \mid f_{\tilde{t}}, \beta, \varphi, \Theta\right)$ , ou seja, gerar  $\left\{ \left( f_{\tilde{t+b}}^{(1)} \right), \dots, \left( f_{\tilde{t+b}}^{(M)} \right) \right\}$ , de modo que

$$\hat{p}\left(y_{\tilde{t+b}} \mid y\right) = M^{-1} \sum_{j=1}^M p\left(y_{\tilde{t+b}} \mid f_{\tilde{t+b}}^{(j)}, \beta^{(j)}, \varphi^{(j)}, \Theta^{(j)}\right)$$

é uma aproximação de Monte Carlo para  $p\left(y_{\tilde{t+b}} \mid y\right)$ . Analogamente, uma amostra  $\left\{ y_{\tilde{t+b}}^{(1)}, \dots, y_{\tilde{t+b}}^{(M)} \right\}$  de  $p\left(y_{\tilde{t+b}} \mid y\right)$  é obtida amostrando  $y_{\tilde{t+b}}^{(j)}$  de  $p\left(y_{\tilde{t+b}} \mid f_{\tilde{t+b}}^{(j)}, \beta^{(j)}, \varphi^{(j)}, \Theta^{(j)}\right)$ , para  $j = 1, \dots, M$ .

## 4 RESULTADOS DA APLICAÇÃO A COMPONENTES SELECIONADOS DA CTBB

Além dos 20 impostos, consideramos que a inclusão de certas variáveis macroeconômicas seria importante na medida em que essas variáveis estão necessariamente associadas à geração desses impostos. Desse modo, após uma seleção parcimoniosa, consideramos incluir o seguinte conjunto de variáveis macroeconômicas:

*PIB*: Produto Interno Bruto deflacionado pelo Índice de Preços ao Consumidor Amplo (IPCA). Fonte: Banco Central do Brasil (BCB);

*DLSP*: Valor mensal do estoque da Dívida Líquida do Setor Público consolidado. Fonte: BCB;

*IPCA*: Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE);

*PIND*: Índice do produto industrial. Fonte: IBGE; e

*CRED*: Crédito privado. Fonte: BCB.

Assim, nossa base de dados é composta por 25 variáveis que foram padronizadas com relação a sua média e desvio-padrão amostral para facilitar a comparação entre elas, o que não afeta o processo de modelagem nem a análise da estrutura fatorial, já que a estrutura de variância-covariância não será afetada. Foram ajustados seis modelos ao todo, utilizando estrutura  $AR(1)$  para os fatores, e variando o número de fatores comuns para  $k = 1, \dots, 6$ .

TABELA 2  
Critério de comparação de modelos

$\rho$	1					
$k$	1	2	3	4	5	6
EQM(p)	31,4467	20,2010	18,0757	17,9364	17,7106	18,4618

Fonte: Elaboração dos autores.

Como pode ser observado na tabela 2, o modelo que apresentou melhor desempenho em termos de previsão fora da amostra de acordo com o critério  $EQM(p)^{12}$  foi aquele com  $p = 1$ ,  $k = 5$ . Para a aproximação das posteriores, foram usadas duas cadeias partindo de pontos diferentes.

No anexo A, são apresentados os resultados do modelo. Por economia, apresentamos os resultados gráficos para apenas um conjunto restrito de variáveis: duas variáveis macroeconômicas (PIB e produção industrial) e cinco impostos (ICMS, II, IRRF – rendimento do trabalho, Cofins, PIS/PASEP e IPI – vinculado à importação). Para cada uma dessas variáveis, apresentamos gráficos para: *i*) o ajustamento do modelo no período janeiro/1995-dezembro/2007; *ii*) as previsões pontuais para o período janeiro/2008-novembro/2008; e, por fim, *iii*) a variabilidade da variável explicada pelos fatores.

Observando os resultados do anexo A, percebe-se inicialmente que as variáveis em questão têm uma forte estrutura sazonal que não pode ser ignorada na análise. Com efeito, nota-se que o método usado para tratar a sazonalidade gerou um alto grau de ajustamento na série de cada variável. Outro resultado interessante desse estudo foi o fato de aparentemente haver uma queda significativa e generalizada na porcentagem da variabilidade das variáveis explicada pelos fatores no período entre 2000 e 2002. Essa característica se repetiu, em geral, para todos os 20 impostos e as 5 variáveis macroeconômicas da nossa base de dados, e nos indica que nesse período houve uma mudança radical na dinâmica das variáveis do estudo, que pararam de responder aos agentes usuais (representados em nosso estudo pelos 5 fatores) como de fato ocorre antes e depois do período 2000-2002. Interpretamos esses resultados como indícios adicionais de que o ajuste da economia brasileira à crise cambial de 1999 representou uma importante “quebra estrutural” em relação ao primeiro quinquênio após o Plano Real (como apontado, entre vários outros, por MENDONÇA, SANTOS e SACHSIDA, 2009)

Por fim, cumpre notar que as previsões obtidas com a nossa modelagem se mostraram superiores àquelas geradas por modelos que são comumente usados na literatura, tais como modelos de regressão univariados com e sem parâmetros variando no tempo.

## 5 COMENTÁRIOS FINAIS

O presente trabalho apresentou a aplicação de um MFD bayesiano aos dados mensais da arrecadação de 20 dos mais importantes (em termos de valor) tributos brasileiros. Isso permitiu que as inter-relações entre as variáveis fossem levadas em consideração na hora da modelagem, o que difere das abordagens recentes usadas para lidar com a CTBB.

Com base em estudos preliminares, acreditamos que as previsões fornecidas por essa modelagem são promissoras e merecem uma comparação formal com outros modelos geralmente usados para modelar a CTBB.

---

12. Erro quadrado médio  $p$  passos à frente;  $EQM(p) = p^{-1} \sum_{i=1}^p \epsilon_{T+i}^2$  onde  $T$  é a última observação usada na estimação.

Além de expandir nossas análises no futuro para comparar resultados de diferentes modelos geralmente usados na literatura, pretende-se incorporar uma estrutura sazonal dinâmica para verificar se haverá melhora no ajuste do modelo, assim como ver se essa estrutura dinâmica irá trazer benefícios no sentido de entender melhor as variações da sazonalidade ao longo do tempo e identificar mudanças estruturais na mesma.

Outro ponto que gostaríamos de abordar posteriormente seria o tópico de intervenções pontuais nesse modelo, de modo que se possam incorporar as mudanças bruscas na economia, que não se refletem nos dados, como é o caso da crise que assolou o sistema financeiro recentemente e que continua gerando consequências de forte impacto em nossas variáveis de interesse.

## REFERÊNCIAS

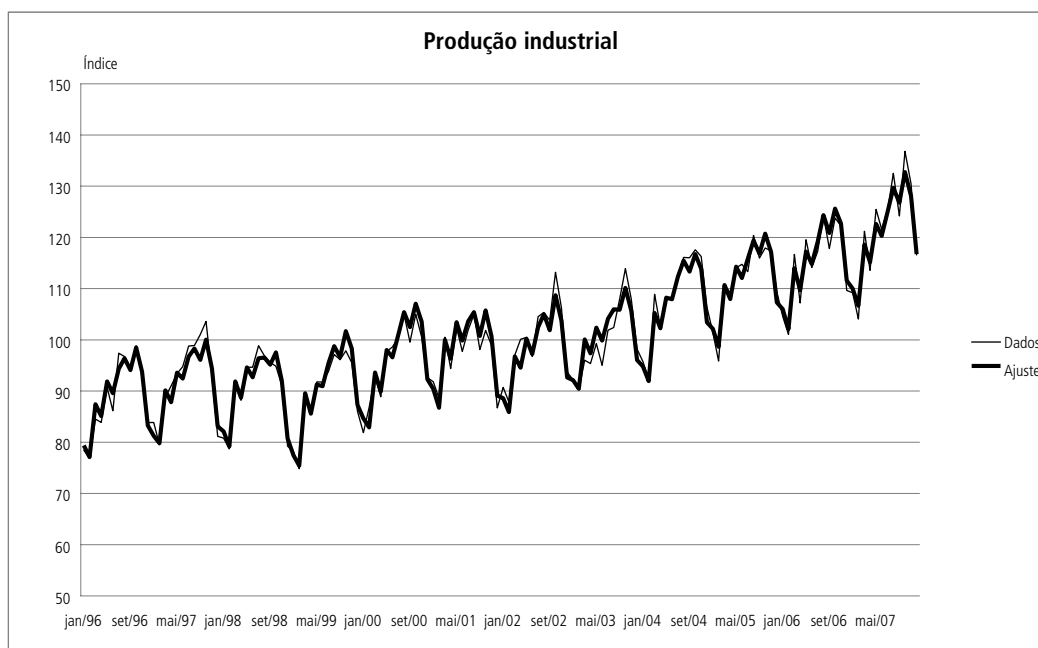
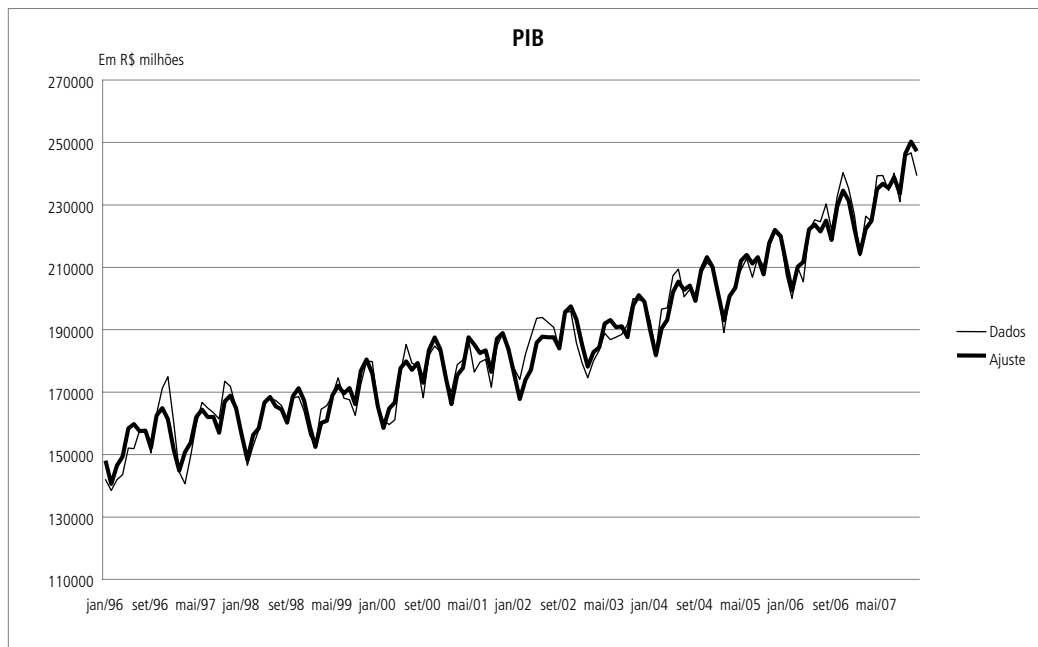
- AGUILAR, O.; WEST, M. Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics*, v. 18, p. 338-357, 2000.
- BARTHOLOMEW, D. J. *Latent variable models and factor analysis*. London: Charles Griffin, 1987.
- CARTER, C. K.; KOHN, R. On Gibbs sampling for state space models. *Biometrika*, v. 81, n. 3, p. 541-553, 1994.
- CARVALHO, J. C. *Tendência de longo prazo das finanças públicas no Brasil*. Ipea, 2001 (Texto para Discussão, n. 795).
- GAMERMAN, D.; LOPES, H. *Markov Chain Monte Carlo*. Chapman & Hall, 2006.
- GEWEKE, J.; ZHUO, G. Measuring the pricing error of the arbitrage pricing theory. *The Review of Financial Studies*, v. 9, p. 557-587, 1996.
- IBGE. *Sistema de Contas Nacionais: Brasil*. Rio de Janeiro: IBGE, 2004 (Série Relatórios Metodológicos, v. 24).
- HAMILTON, J. *Time series analysis*. Princeton University Press, 1993.
- HERNÁNDEZ, B. *Um modelo econométrico da conta corrente do governo no Brasil – 1951/95*. Ipea, 1998 (Texto para Discussão, n. 543).
- KIM, C. J.; NELSON, C. R. *State-space models with regime switching*. Massachusetts: The MIT Press, 1999.
- MELO, B. *Modelo de previsão para arrecadação tributária*. VI Prêmio Tesouro Nacional, 2001.
- MENDONÇA, M. J. C.; SANTOS, C. H. M. dos; SACHSIDA, A. *Revisitando a função de reação fiscal no Brasil pós-real: uma abordagem de mudanças de regime*. *Estudos Econômicos*, v. 39, n. 4, 2009.
- MENDONÇA, M. J. C.; PIRES, M. C.; MEDRANO, L. T. *Administração e sustentabilidade da dívida pública no Brasil: uma análise para o período 1996-2007*. Ipea, 2008 (Texto para Discussão, n. 1.342, a ser publicado na revista *Pesquisa e Planejamento Econômico*).

- PORTUGAL, C.; PORTUGAL, M. Os efeitos da inflação sobre o orçamento do governo: uma análise empírica. *Estudos Econômicos*, v. 31, n. 2, p. 239-283, 2001.
- REZENDE, F. R.; OLIVEIRA, F.; ARAÚJO, E. *O dilema fiscal: remendar ou reformar?* Confederação Nacional da Indústria e Fundação Getúlio Vargas, 2007.
- SÁFADI, T.; PENA, D. Bayesian analysis of dynamic factor models: an application to air pollution and mortality in São Paulo, Brazil. *Environmetrics*, v. 19, n. 6, p. 582-601, 2007.
- SANTOS, C. H. M. dos; RIBEIRO, M. B.; GOBETTI, S. W. *A evolução da carga tributária bruta brasileira no período 1995-2007: tamanho, composição e especificações econométricas agregadas*. Ipea, 2008 (Texto para Discussão, n. 1.350).
- SIQUEIRA, M. *Melhorando a previsão de arrecadação tributária federal através da utilização de modelos de séries temporais*. VII Prêmio Tesouro Nacional, 2002.
- WEST, M.; HARRISON, J. *Bayesian forecasting and dynamic models*. Springer, 1997.

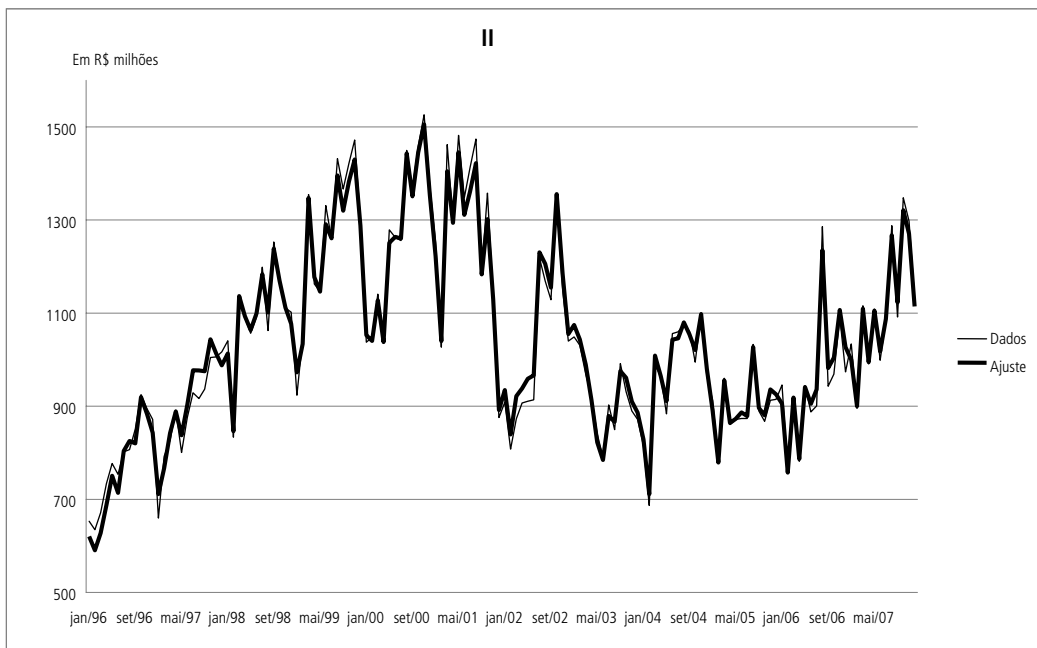
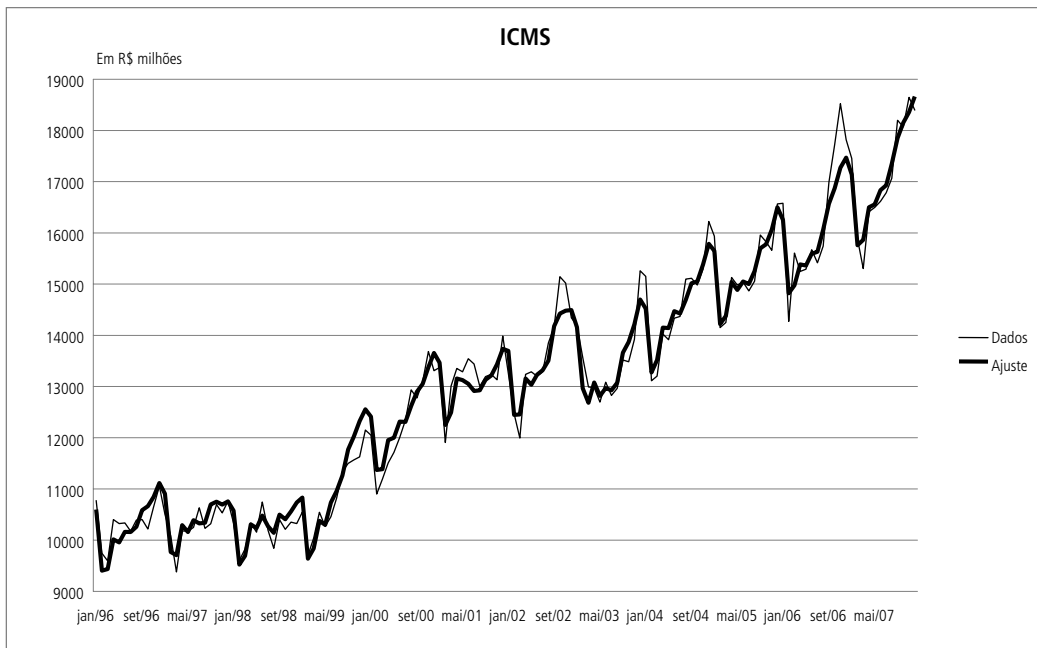
## ANEXO A

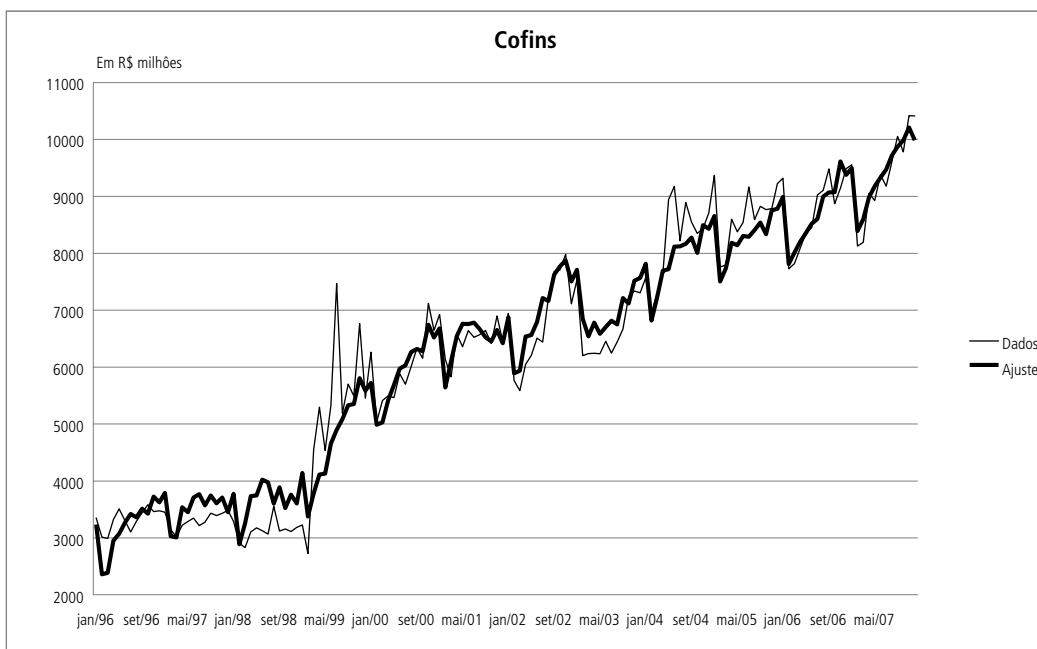
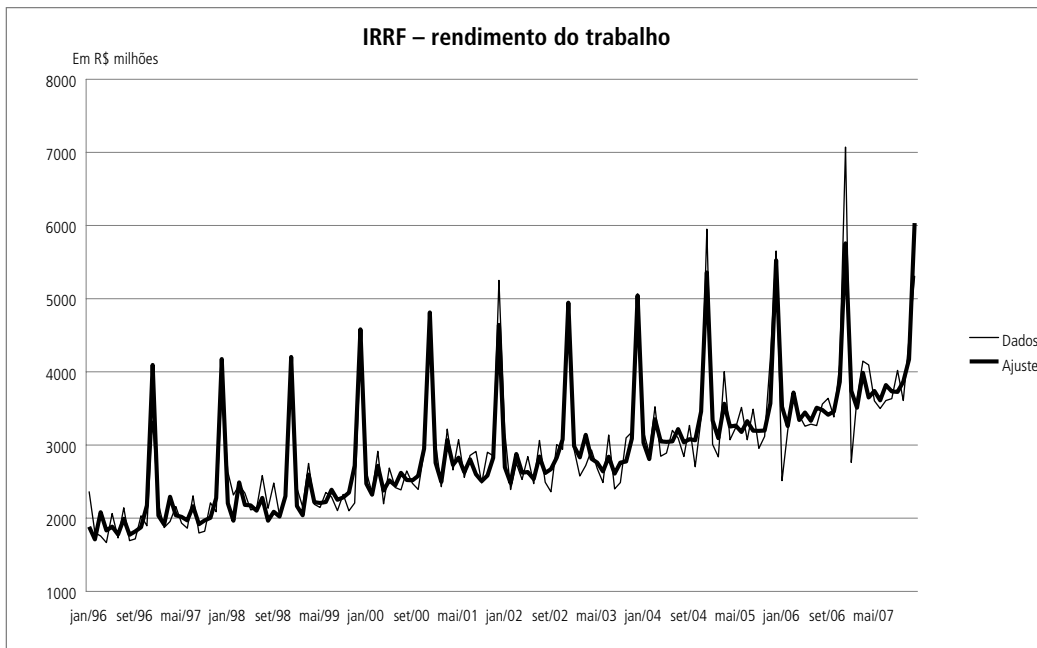
GRÁFICO A.1

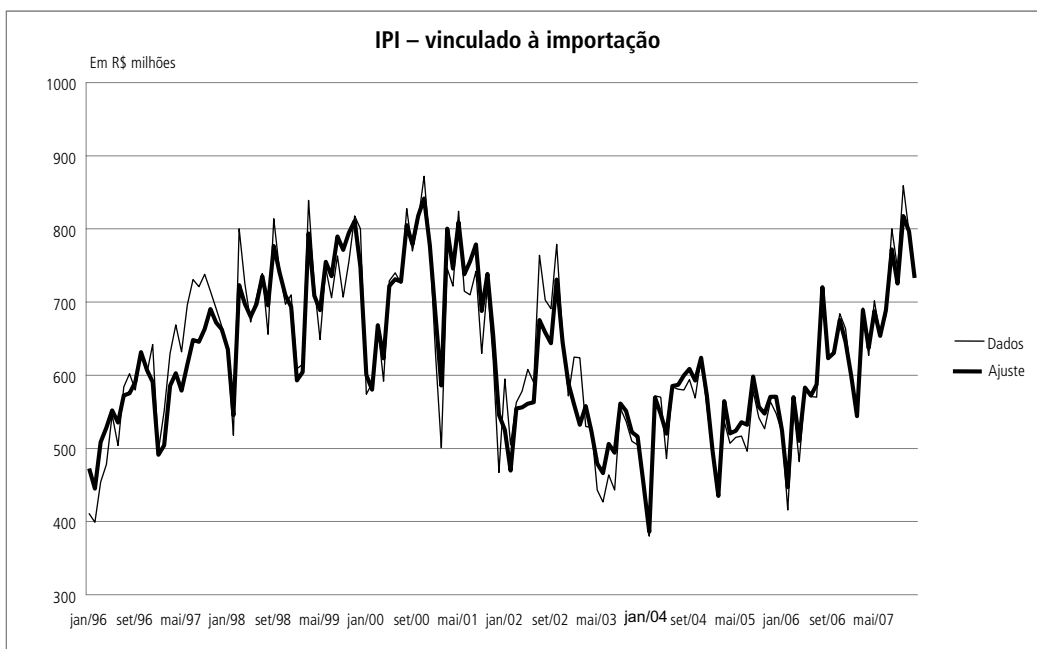
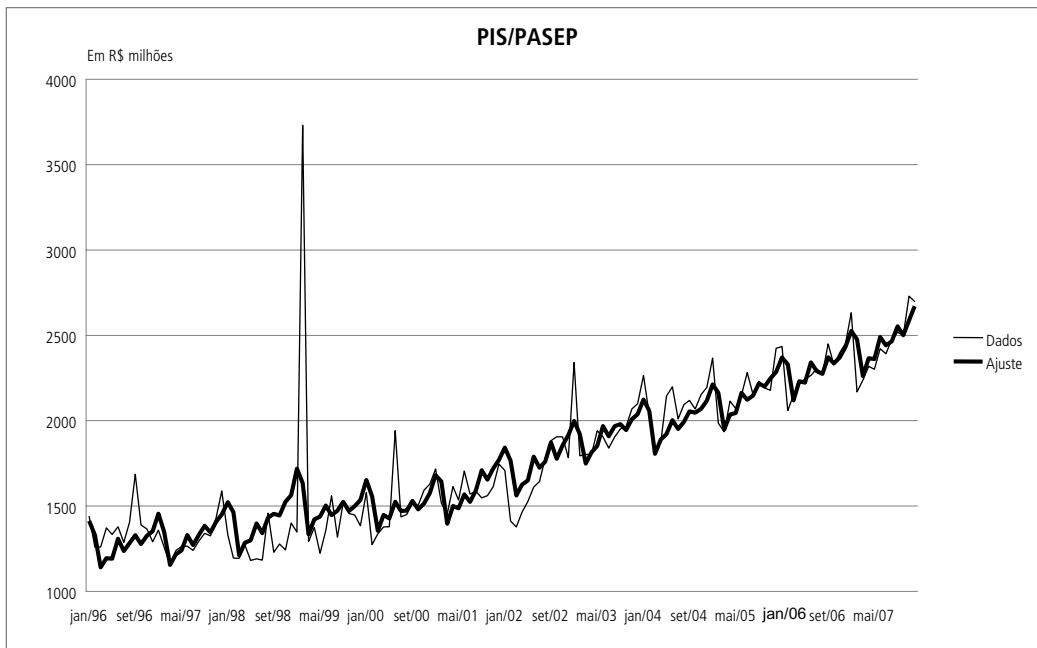
Variáveis PIB, produção industrial, ICMS, II, IRRF – rendimento do trabalho, Cofins, PIS/PASEP e IPI – vinculado à importação no período janeiro/1996-dezembro/2007: ajuste do modelo<sup>13</sup>



13. Com exceção do índice de produção industrial, as demais variáveis estão expressas em valores reais de dezembro de 2007.





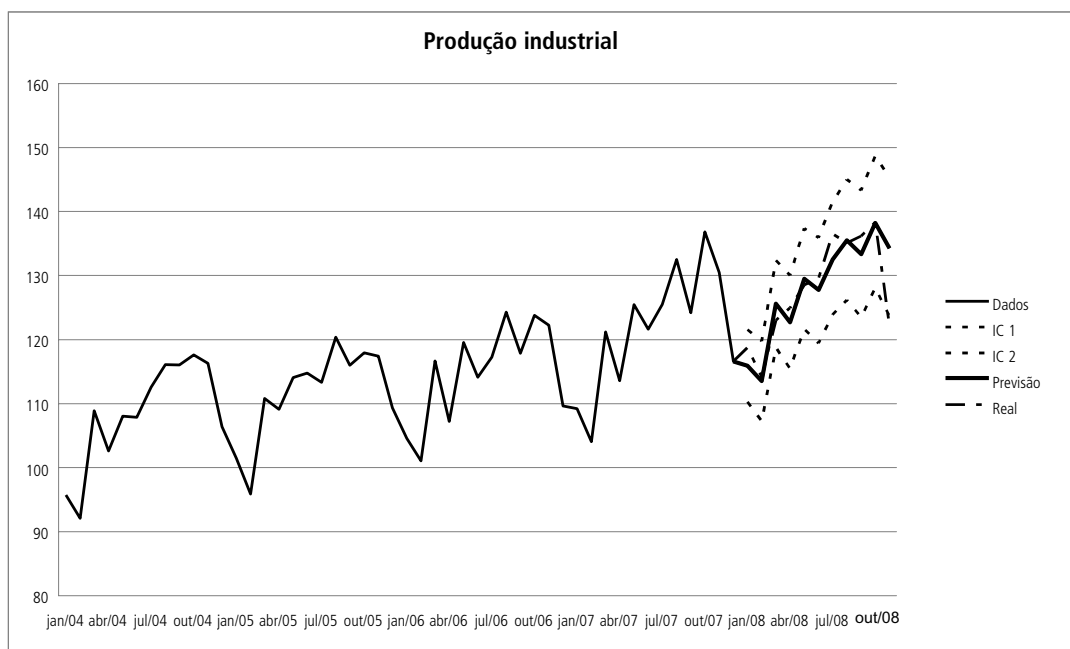
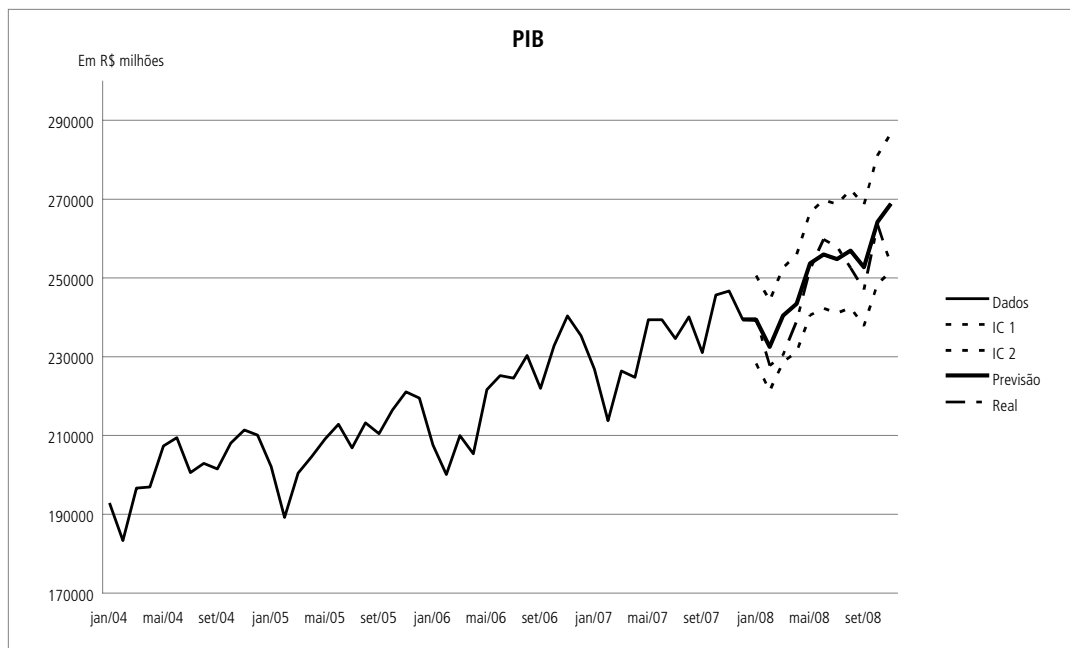


Fonte: Elaboração dos autores.

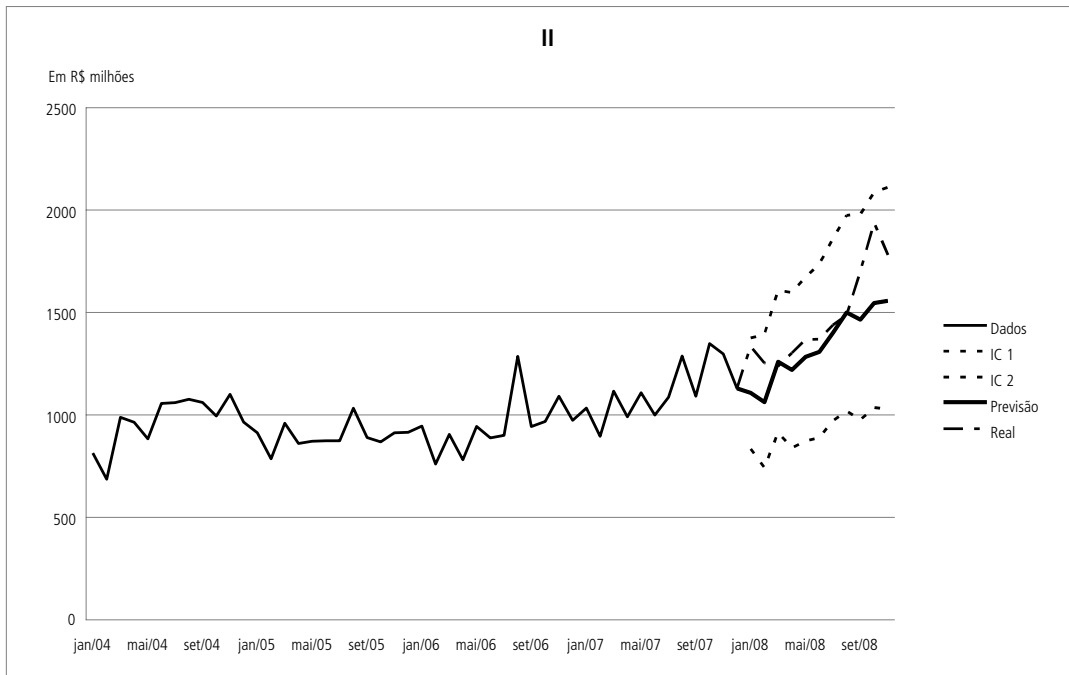
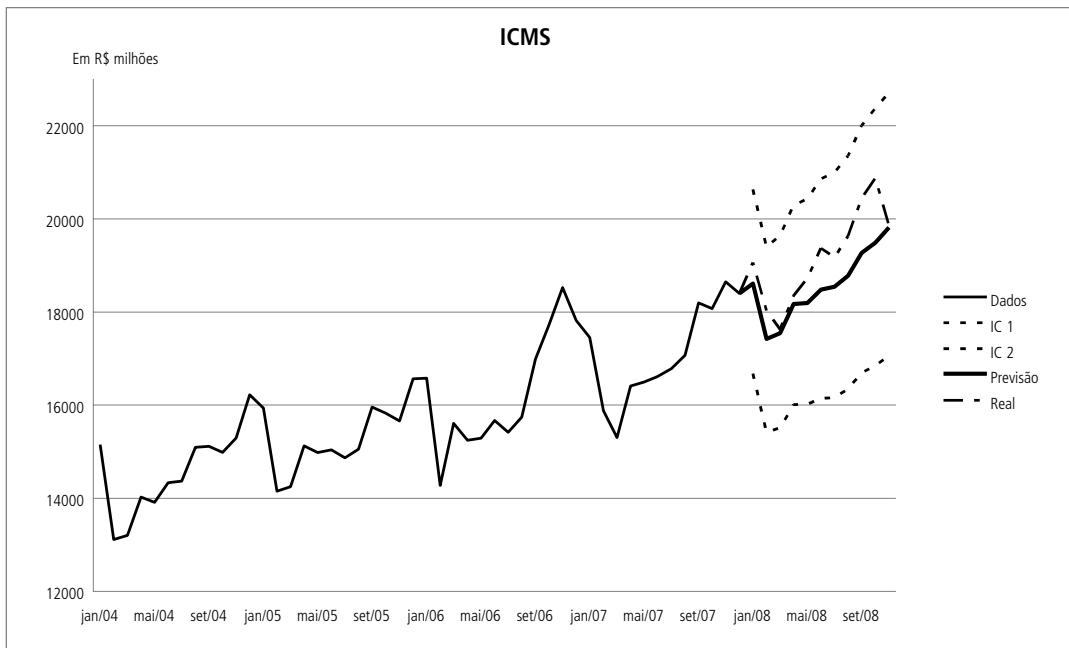


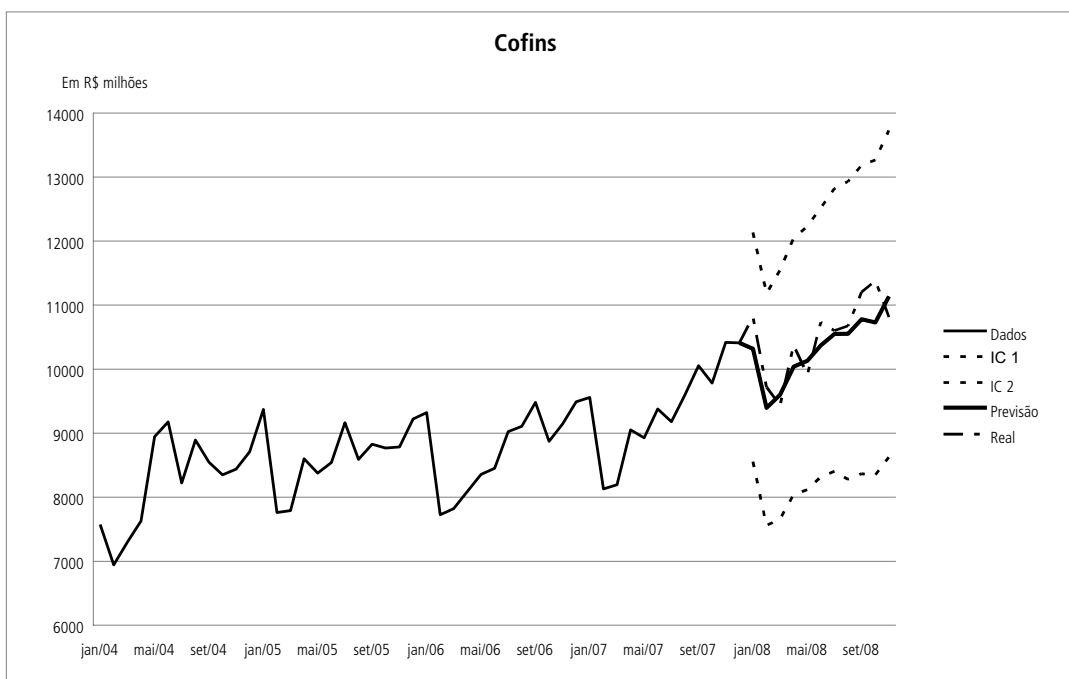
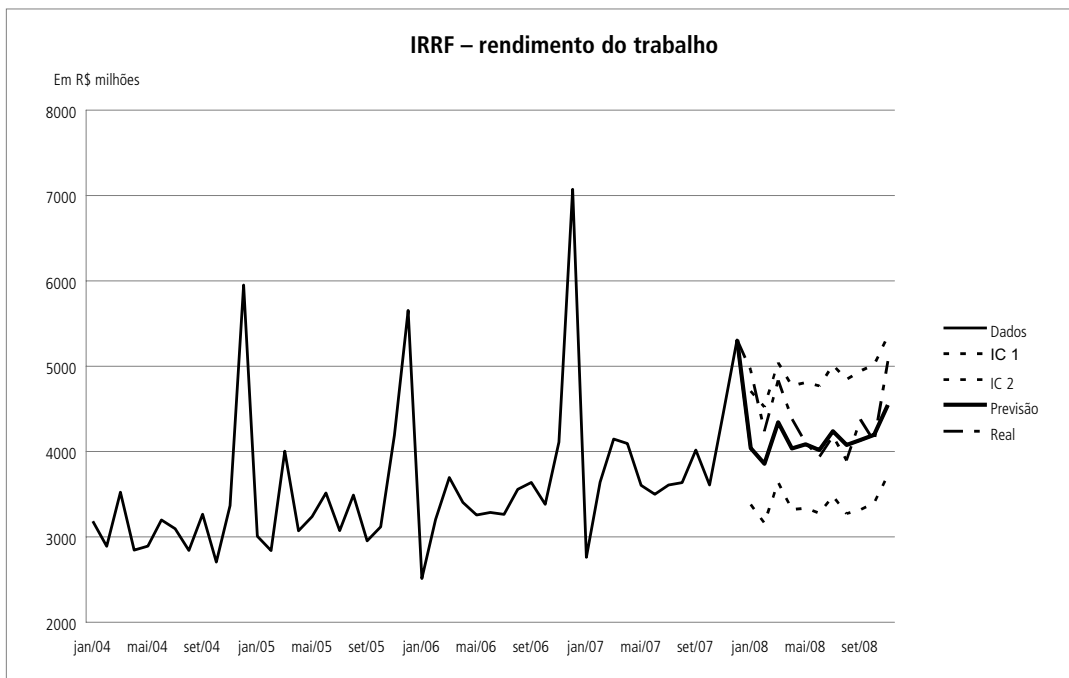
GRÁFICO A.2

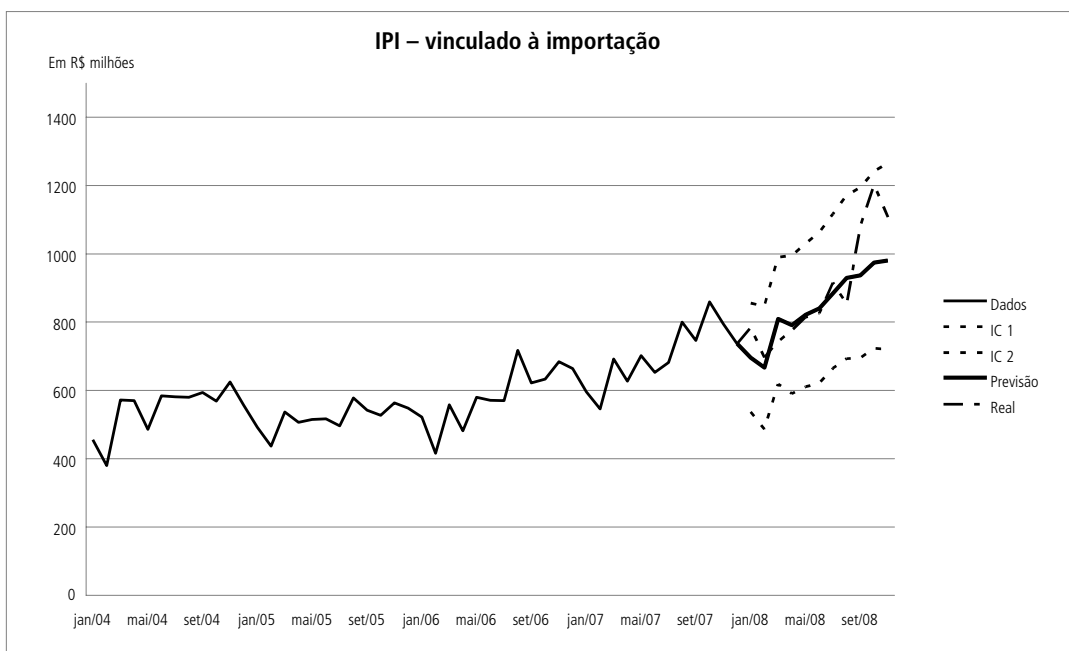
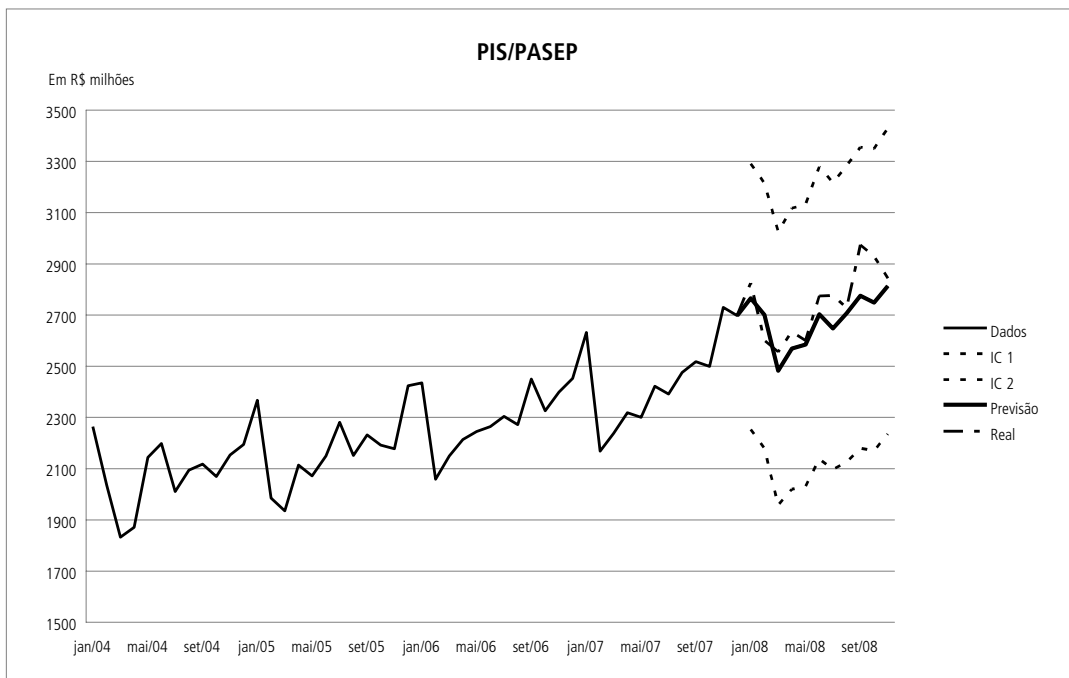
**Variáveis PIB, produção industrial, ICMS, II, IRRF-rendimento do trabalho, Cofins, PIS/PASEP e IPI – vinculado à importação no período janeiro/2004 – dezembro/2007<sup>14</sup>**



14. Com exceção do índice de produção industrial, as demais variáveis estão expressas em valores reais de dezembro de 2007.



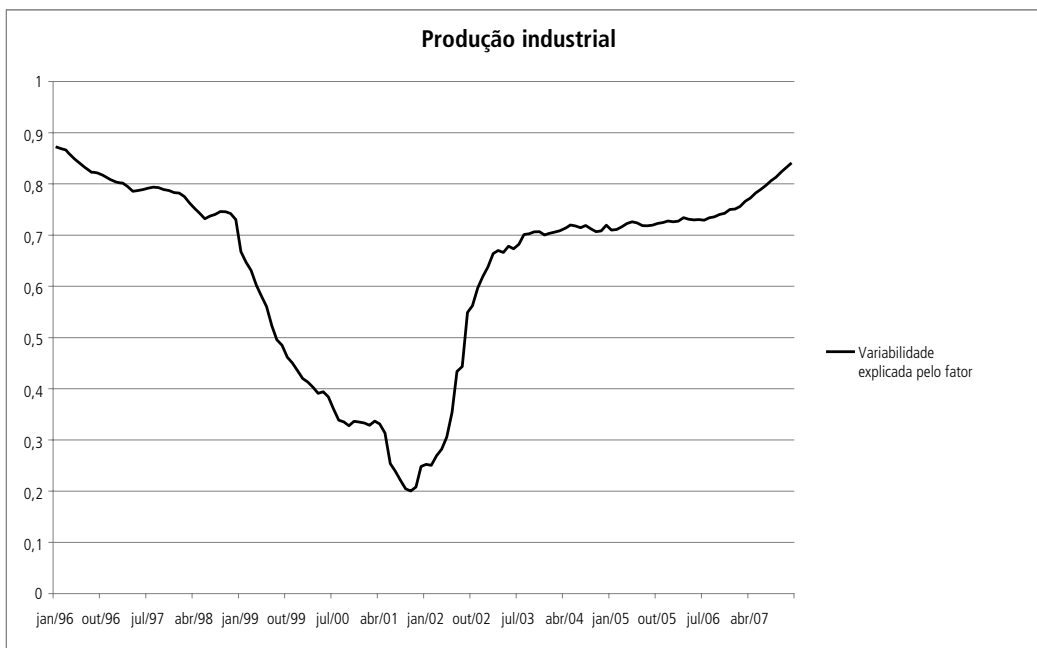
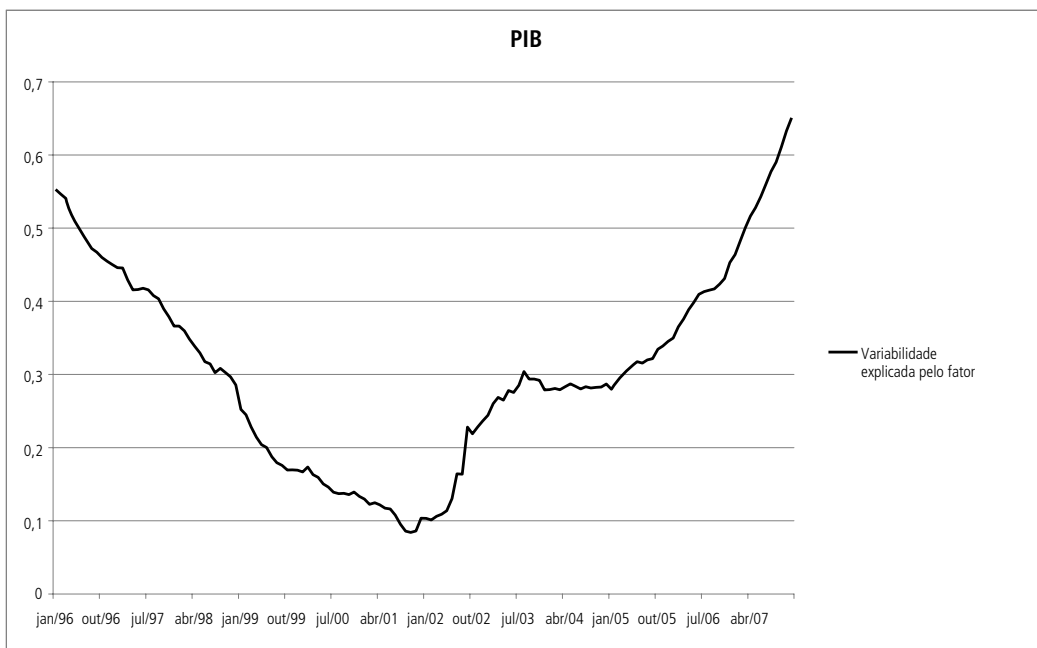


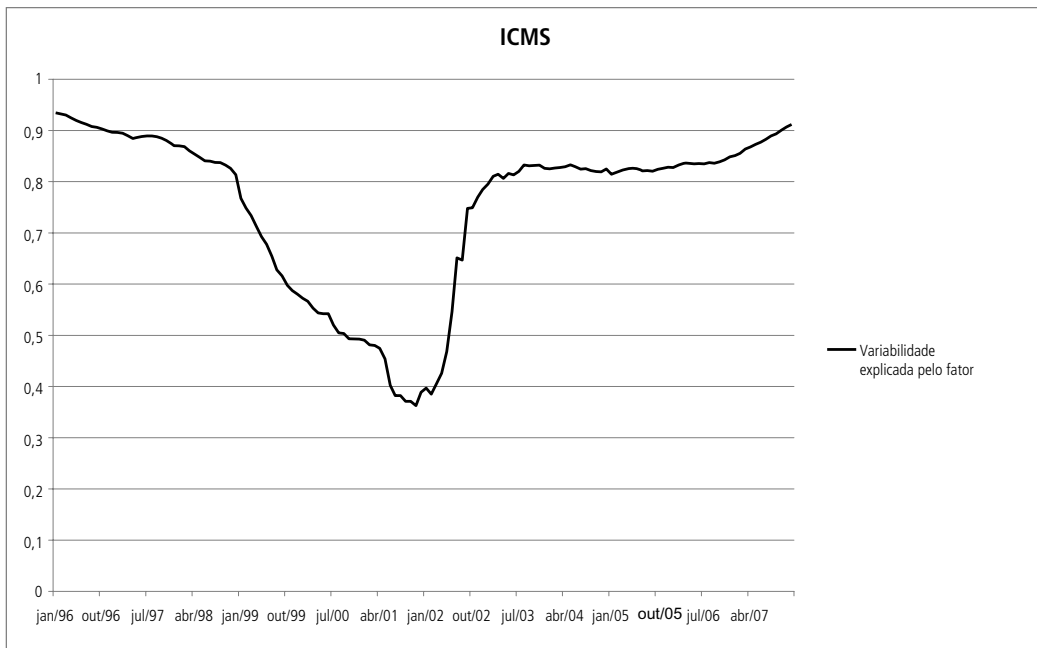


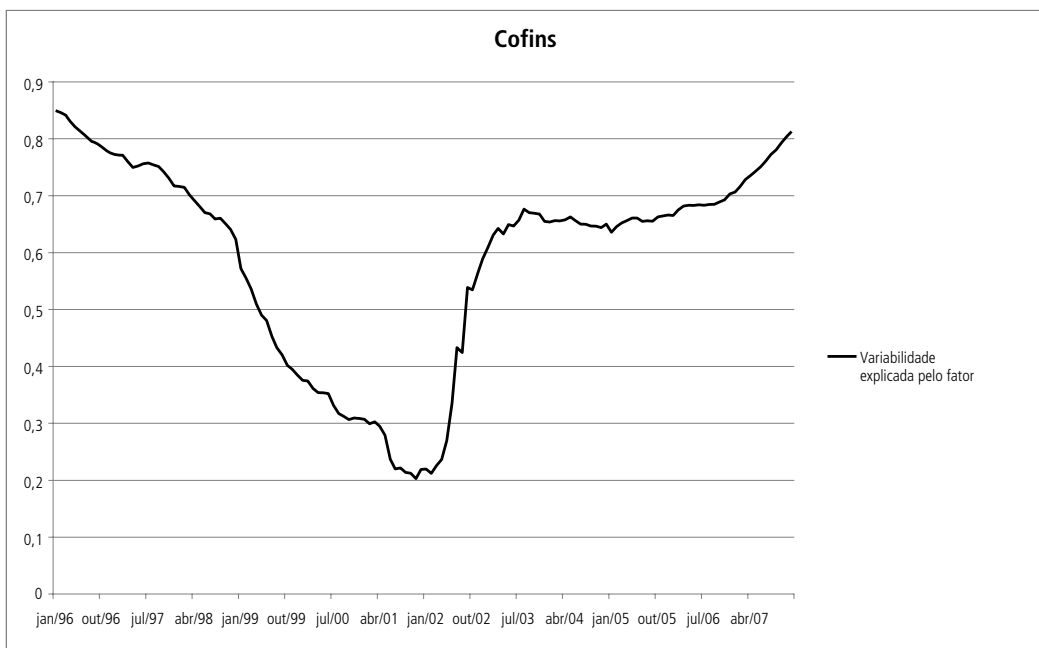
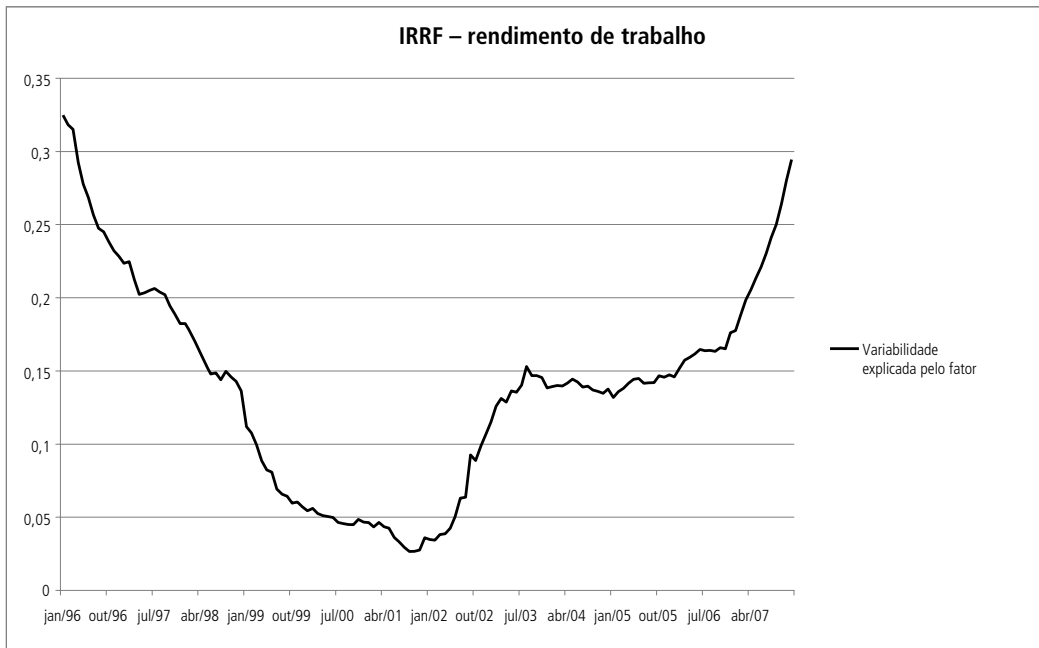
Fonte: Elaboração dos autores.

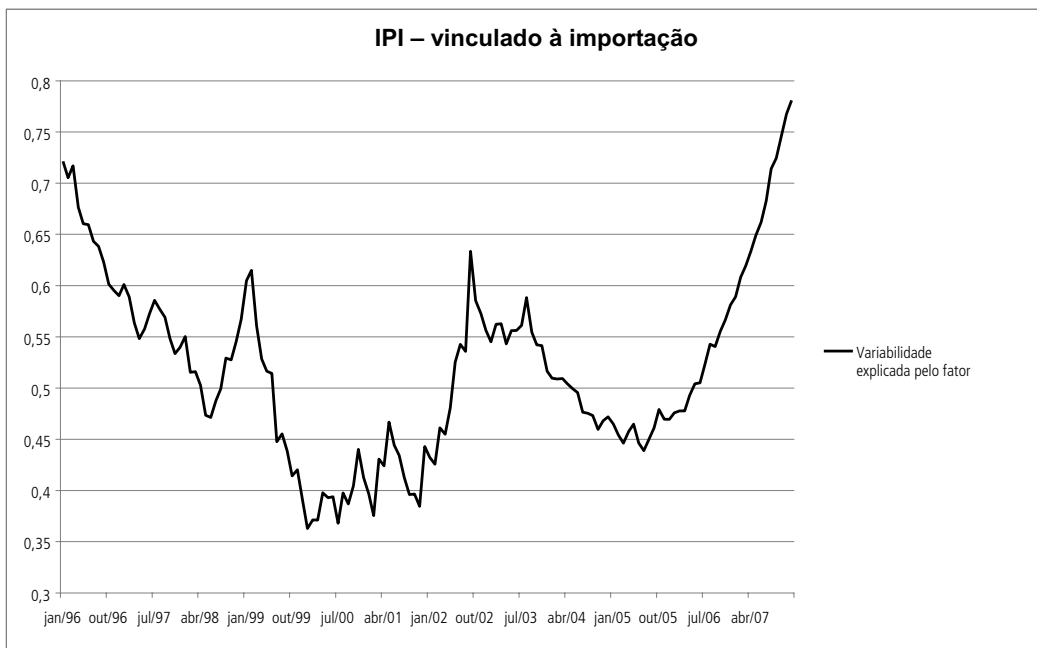
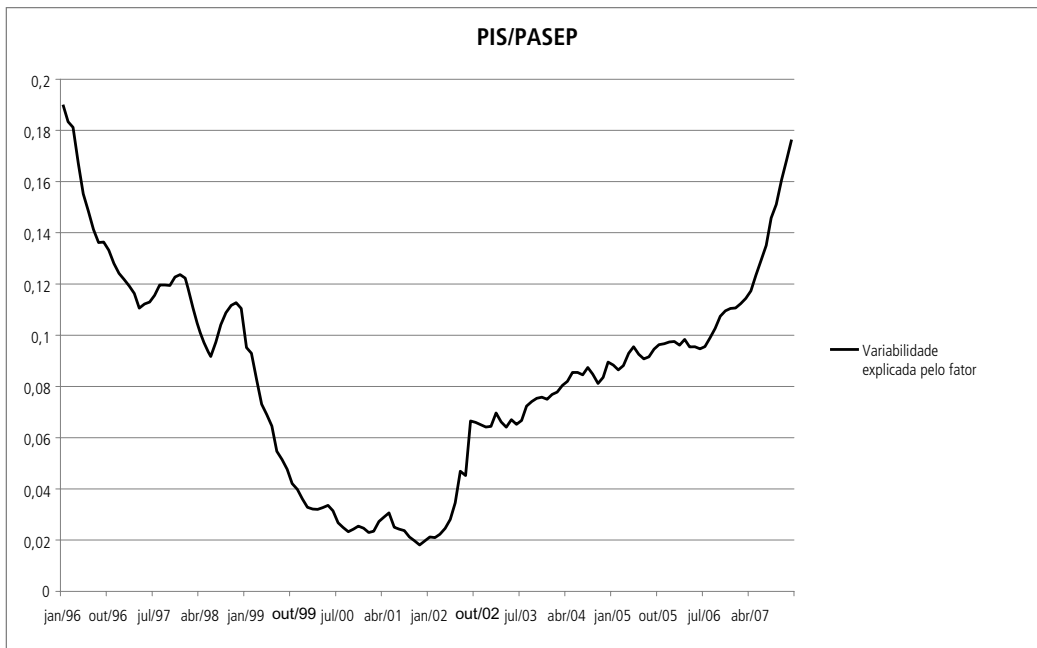
GRÁFICO A.3

**Porcentagem da variabilidade das variáveis PIB, produção industrial, ICMS, II, IRRF–rendimento do trabalho, Cofins, PIS/PASEP e IPI–vinculado à importação que é explicada pelos fatores ao longo do tempo**









Fonte: Elaboração dos autores.



## ANEXO B

### 1 DESCRIÇÃO DO ALGORITMO MCMC PARA O MODELO FATORIAL DINÂMICO

Podemos representar o modelo (1) em uma estrutura de espaço de estados, do seguinte modo:

$$\begin{aligned} (y_t - \phi X_t) &= F^E \tilde{f}_t^E + \underline{e}_t, \quad \underline{e}_t \sim N(0, \Sigma) \\ \tilde{f}_t^E &= G^E \tilde{f}_{t-1}^E + \underline{w}_t^E, \quad \underline{w}_t^E \sim N(0, Q^E) \end{aligned} \quad (2)$$

onde:

$$G_{pkxp}^E = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{p-1} & \Gamma_p \\ I_{kxk} & \underline{0}_{kxk} & \cdots & \underline{0}_{kxk} & \underline{0}_{kxk} \\ \underline{0}_{kxk} & I_{kxk} & \cdots & \underline{0}_{kxk} & \underline{0}_{kxk} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \underline{0}_{kxk} & \underline{0}_{kxk} & \cdots & I_{kxk} & \underline{0}_{kxk} \end{bmatrix}, \quad \tilde{f}_t^E = \begin{bmatrix} f_t \\ \tilde{f}_{t-1} \\ \vdots \\ \tilde{f}_{t-p+1} \end{bmatrix}, \quad \underline{w}_t^E = \begin{bmatrix} \underline{w}_t \\ \underline{0}_{kx1} \\ \vdots \\ \underline{0}_{kx1} \end{bmatrix},$$

$$Q_{pkxp}^E = \begin{bmatrix} \Lambda_{kxk} & \underline{0}_{kxk(p-1)} \\ \underline{0}_{k(p-1)xk} & \underline{0}_{k(p-1)xk(p-1)} \end{bmatrix}, \quad F_{mxpk}^E = [\beta_{mxk} \quad \underline{0}_{mxk} \quad \underline{0}_{mxk} \quad \cdots \quad \underline{0}_{mxk}]$$

### 2 POSTERIORIS COMPLETAS NECESSÁRIAS PARA A IMPLEMENTAÇÃO DO AMOSTRADOR DE GIBBS

Para  $i = 1, \dots, k$ , escreva  $\underline{\beta}_i = (\beta_{i1}, \dots, \beta_{ii})^T$ , que são os elementos diferentes de zero. Para os restantes das linhas  $i = k+1, \dots, m$ , escreva  $\underline{\beta}_i = (\beta_{i1}, \dots, \beta_{ik})$ . Similarmente, para  $i = 1, \dots, k$  denote  $F_i$ , a matriz  $Txi$ , contendo as primeiras  $i$  colunas de  $F$ , e para todo  $i$  seja  $y_i$  a  $i$ -ésima coluna de  $y$ ,  $S_i$  a  $i$ -ésima coluna de  $(X\phi^T)$  com  $X = (X_1^T, \dots, X_T^T)^T$  e  $y_i^* = y_i - S_i$ . Seja  $\underline{\gamma}_j = (\gamma_{1j}, \gamma_{2j}, \dots, \gamma_{pj})^T$  o vetor contendo o  $j$ -ésimo elemento da diagonal para cada uma das  $p$  matrizes  $\Gamma_1, \Gamma_2, \dots, \Gamma_p$ . Defina também para  $j = 1, \dots, k$

$$\mathfrak{S}_j = \begin{pmatrix} f_{p,j} & f_{p-1,j} & \cdots & f_{1,j} \\ f_{p+1,j} & f_{p,j} & \cdots & f_{2,j} \\ \vdots & \vdots & \vdots & \vdots \\ f_{T-1,j} & f_{T-2,j} & \cdots & f_{T-p,j} \end{pmatrix} \text{ e } H_j = \begin{pmatrix} f_{p+1,j} \\ f_{p+2,j} \\ \vdots \\ f_{T,j} \end{pmatrix}$$

- Para amostrar da condicional completa de  $f_t$ , siga os passos a seguir:

1) Aplicar filtro de Kalman para calcular  $m_t = E[f_t^E | D_t]$  e  $C_t = V[f_t^E | D_t]$  para  $t = 1, \dots, T$ , do seguinte modo:

$$a_t = G^E m_{t-1}, \quad R_t = G^E C_{t-1} G^{E^T} + Q^E$$

$$f_t = F^E a_t, \quad Q_t = F^E R_t F^{E^T} + \Sigma$$

$$A_t = R_t F^{E^T} Q_t^{-1}, \quad e_t = (y_t - S_t) - f_t$$

e armazenar  $m_t$  e  $C_t$ ,  $t = 1, \dots, T$ . No final, use  $m_t$  e  $C_t$  para gerar  $f_t^E | \cdot \sim N(m_t, C_t)$ .

2) Defina  $f_t^{E^*}$  como sendo os primeiros  $k$  elementos de  $f_t^E$ ,  $G^{E^*}$  como sendo as primeiras  $k$  linhas de  $G^E$  e  $Q^{E^*}$  como sendo o primeiro bloco  $k \times k$  da matriz  $Q^E$ . Para  $t = T-1, T-2, \dots, 1$  faça:

$$m_t^* = m_t + C_t G^{E^*T} \left( G^{E^*} C_t G^{E^*T} + Q^{E^*} \right)^{-1} \left( f_{t+1}^{E^*} - G^{E^*} m_t \right),$$

$$C_t^* = C_t - C_t G^{E^*T} \left( G^{E^*} C_t G^{E^*T} + Q^{E^*} \right)^{-1} G^{E^*} C_t$$

e gere  $f_t^E | \cdot, f_{t+1}^{E^*} \sim N(m_t^*, C_t^*)$  e mantenha somente os  $k$  primeiros elementos gerados dessa forma. Esses primeiros  $k$  elementos formam uma amostra de  $f_t$  da condicional completa.

- A posteriori condicional completa de  $\Sigma$  se reduz a um conjunto de  $m$  gammas inversas independentes, com  $\sigma_i^2 \sim IG\left(\frac{n_\sigma^*}{2}, \frac{n_\sigma^* S_\sigma^*}{2}\right)$ , onde:

$$n_\sigma^* = n_\sigma + T$$

$$n_\sigma^* S_\sigma^* = n_\sigma S_\sigma + \left( y_i^* - F \beta_i^* \right)^T \left( y_i^* - F \beta_i^* \right)$$

- A posteriori condicional completa de  $\beta$  divide-se em:

para  $i = 1, \dots, k$ ,  $\beta_i \sim N(\mu_i, V_i) 1(\beta_{ii} = 1)$  onde  $\mu_i = V_i^{-1} (V_0^{-1} \mu_0 1_i + \sigma_i^{-2} F_i^T y_i^*)$  e  $V_i^{-1} = V_0^{-1} I_i + \sigma_i^{-2} F_i^T F_i$ ;

para  $i = k+1, \dots, m$ ,  $\beta_i \sim N(\underline{\mu}_i, V_i)$  onde  $\underline{\mu}_i = V_i(V_0^{-1}\mu_0 \mathbf{1}_k + \sigma_i^{-2} F^T y_i^*)$  e  $V_i^{-1} = V_0^{-1} I_k + \sigma_i^{-2} F^T F$

- A posteriori condicional completa de  $\gamma_j$  é dada por:

para  $j = 1, \dots, k$ ,  $\gamma_j \sim N(A_j, B_j)$  onde  $A_j = B_j(S_r^{-1} m_r \mathbf{1}_p + \lambda_j^{-1} \mathfrak{F}_j^T H_j)$  e  $B_j = S_r^{-1} I_{p \times p} + \lambda_j^{-1} \mathfrak{F}_j^T \mathfrak{F}_j$

- A posteriori condicional completa de  $\Lambda$  se reduz a um conjunto de  $k$  gammas inversas independentes, com  $\lambda_j \sim IG\left(\frac{n_\lambda^*}{2}, \frac{n_\lambda^* S_\lambda^*}{2}\right)$ , onde:

$$n_\lambda^* = n_\lambda + T$$

$$n_\lambda^* S_\lambda^* = n_\lambda S_\lambda + (H_j - \mathfrak{F}_j \gamma_j)^T (H_j - \mathfrak{F}_j \gamma_j)$$

- A posteriori condicional completa de  $\varphi_i$  é dada por:

para  $i = 1, \dots, m$ ,  $\varphi_i \sim N(C_i, D_i)$  onde  $C_i = D_i(C_s^{-1} m_s \mathbf{1}_{12} + \sigma_i^{-1} X(y_i - F\beta_i))$  e  $D_i = C_s^{-1} I_{12 \times 12} + \sigma_i^{-2} X^T X$

### 3 PREVISÃO DO MODELO FATORIAL DINÂMICO COM COMPONENTE SAZONAL COM $p > 1$

Passa o modelo (1) para a estrutura de espaço de estados (2) e utilize a mesma metodologia apresentada na seção 3.5, substituindo  $f_t$ ,  $\Lambda$ ,  $\beta$  e  $\Gamma_1$  por  $f_t^E$ ,  $Q^E$ ,  $F^E$  e  $G^E$  respectivamente.



© Instituto de Pesquisa Econômica Aplicada – Ipea 2009

**EDITORIAL**

**Coordenação**

Iranilde Rego

**Supervisão**

Andrea Bossle de Abreu

**Revisão**

Lucia Duarte Moreira

Eliezer Moreira

Elisabete de Carvalho Soares

Fabiana da Silva Matos

Gilson Baptista Soares

Míriam Nunes da Fonseca

Roberta da Costa de Sousa

**Editoração**

Roberto das Chagas Campos

Aeromilson Mesquita

Camila Guimarães Simas

Carlos Henrique Santos Vianna

Aline Cristine Torres da Silva Martins (estagiária)

**Livraria**

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo

70076-900 – Brasília – DF

Fone: (61) 3315-5336

Correio eletrônico: [livraria@ipea.gov.br](mailto:livraria@ipea.gov.br)

Tiragem: 130 exemplares